AD_____

Award Number: W81XWH-06-2-0072

TITLE: Advanced Processing for Biomedical Informatics (APBI)

PRINCIPAL INVESTIGATOR: Craig D. Shriver, M.D.

CONTRACTING ORGANIZATION: Henry M. Jackson Foundation for the Advancement of Military Medicine
Rockville, MD 20852

REPORT DATE: October 2009

TYPE OF REPORT: Addendum to Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(31-072006)*	2. REPORT TYPE	3. DATES COVERED
October 01, 2009	Addendum to Final	*27 SEP 2008 - 26 SEP 2009*

4. TITLE AND SUBTITLE

Advanced Processing for Biomedical Informatics (APBI)

5a. CONTRACT NUMBER
W81XWH-06-02-0072

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Dr. Craig D. Shriver
Email: craig.shriver@na.amedd.army.mil

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

The Henry M. Jackson Foundation for the
Advancement of Military Medicine, Inc.
1401 Rockville Pike, Suite 600
Rockville, MD 20852

8. PERFORMING ORGANIZATION REPORT

Cost Center Number
105134

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Commander, U.S. Army Medical Research and
Material Command, ATTN MCMR-ZC-I
504 Scott Street
Fort Detrick, MD 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)
USAMRAA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

A research consortium including Windber Research Institute (WRI), the US Army Space and Missile Defense Command (USASMDC), the Joyce Murtha Breast Care Center (JMBCC), and the Clinical Breast Care Project (CBCP) Walter Reed Army Medical Center (WRAMC) has been formed to evaluate the use of minimally-invasive methods for screening including mammography, ultrasound, proteomics and genomics, in the serum and breast for early detection of markers for risk of disease or early presence of disease and to facilitate early intervention in medical treatment or lifestyle. The approach focuses on the continuing development/aging that the female breast undergoes through life and its potential sensitivity to environmental and lifestyle factors, particularly as they interact with specific genetic factors.

15. SUBJECT TERMS
Proteomics, Biomedical Informatics, Genomics, Patient Focused, Radiology Information System (RIS), Decision Support System, Cancer Detection, Anomaly Detection Algorithm

16. SECURITY CLASSIFICATION OF: Unclassified			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
			72		*202-782-3416*

COL Craig D. Shriver

202-782-3418

Table of Contents

Advanced Processing for Biomedical Informatics (APBI)
Annual Report

I. Introduction

Currently, mammography is the most widely used technology for breast cancer screening and detection. The final diagnosis is typically made through biopsy of the lesion identified by mammogram reading, followed by pathologic analysis which is the current 'gold standard' for breast cancer diagnosis. When held against this 'gold standard', the reading of mammograms reports a high false positive rate and high false negative rate. Reducing the false discovery rate by mammography can dramatically reduce the number of unnecessary surgeries (false positive) and catch otherwise missed cancer cases (false negative).

The US Army Space and Missile Defense Command has developed algorithms to identify incoming missiles from complicated background signals. Given the similarity of 'target identification' in missile detection and breast cancer identification, a pilot study was proposed for automatic breast cancer identification based on mammograms in combination with other sources of information. This joint study involves the US Army Space and Missile Defense Command, the Walter Reed Army Medical Center, and the Windber Research Institute. Mammograms (digitized) are the centerpiece of the object of study, complemented by ultrasound images, microarray gene expression data, as well the 800 fields of the Core Questionnaire and Pathology Checklist data.

According to the original plan, subjects for this study are CBCP participants enrolled from the Walter Reed Army Medical Centers. As the first step, mammograms from hundreds to thousands of CBCP subjects are used to evaluate the potential for enhanced mammogram image processing algorithms adapted from the SMDC algorithm to improve breast cancer detection. Next, available ultrasound images are used to evaluate the potential of the algorithm adapted from the SMDC algorithms to improve breast cancer detection. Then, the data from these two modalities are fused by applying the Response Surface Methodology models. Furthermore, based on the available resources the whole blood gene expression data from the specimens of the matching subjects are integrated into the data fusion model, which is further combined with the clinical and pathologic data to increase the confidence of breast cancer detection. The end product of this pilot study is a Decision Support System for improved diagnosis of breast cancer.

The SMDC obtained an ACRIN digital mammogram data set for this study consisting of 11,528 mammogram images from 2,467 patients. 1,503 of the images were from the 305 patients who tested positive for breast cancer. This data set was used to develop detection and discrimination algorithms.

The WRI team managed to execute a revised version of the molecular study plan, and complete the data analysis.

II. Body

II.1. Subject selection

The original subject inclusion criteria are:
– Consented to the CBCP Tissue/Blood protocol. (Signed consent on file)
– Complete Questionnaire and Path Checklist on file.
– Tissue /Blood samples available.
– Mammography history/films located at WRAMC.
- De-Identified Mammography films digitized and available for transfer to SMDC.
– Consented to the ICAD study.
– De-Identified Ultrasound study data available for transfer to SMDC.

Throughout the project execution period, WRAMC had sent digitized mammograms or ultrasound images from 208 subjects (114 subjects with ultrasound images, and 135 subjects with digitized mammograms), to SMDC for analysis. A reality check reveals only 41 subjects meeting the above criteria. So for molecular studies, we started from the 208 subjects in the hope that some of the missing data could be available at a later time at which time the molecular study data will be merged with the study of other modality. We went ahead to complete the planned gene expression experiments and data analysis.

II.2. Specimen selection

The available Core Questionnaires and Pathology Checklists were also sent to SMDC to assist in image analysis. We decided to perform gene expression analysis of the samples of these subjects.

We identified that a total of 179 subjects have whole blood PAXgene samples, of which 136 samples satisfy the criteria that each specimen has pathology information, and that the sample came before any major procedure on the subject, some samples have been consumed in previous studies using an older GeneChip from Affymetrix. Finally we were able to conduct experiments on 92 blood samples, with subjects' information shown below.

PARTICIPANT_ID	CATEGORY	SEX	ETHNICITY	MENOPAUSAL_STATUS
100000143	Atypical	Female	African American	Post-menopausal
100001082	Atypical	Female	African American	Post-menopausal
100001651	Atypical	Female	Hispanic (white)	Pre-menopausal
100001663	Atypical	Female	African American	Pre-menopausal
100001705	Atypical	Female	White	Pre-menopausal
100001859	Atypical	Female	White	Post-menopausal
528	Benign	Female	White	Post-menopausal
894	Benign	Female	African American	Pre-menopausal
1083	Benign	Female	White	Post-menopausal
100000144	Benign	Female	White	Pre-menopausal
100000156	Benign	Female	White	Pre-menopausal
100000404	Benign	Female	White	Pre-menopausal
100001047	Benign	Female	White	Surgically menopausal

100001053	Benign	Male	Other	
100001072	Benign	Female	White	Surgically menopausal
100001087	Benign	Female	White	Pre-menopausal
100001090	Benign	Female	Hispanic (other)	Pre-menopausal
100001101	Benign	Female	Hispanic (white)	Post-menopausal
100001102	Benign	Male	African American	
100001163	Benign	Female	Hispanic (white)	Pre-menopausal
100001164	Benign	Female	White	Surgically menopausal
100001232	Benign	Female	White	Pre-menopausal
100001236	Benign	Female	White	Status post hysterectomy
100001249	Benign	Female	White	Pre-menopausal
100001250	Benign	Female	African American	Pre-menopausal
100001253	Benign	Male	White	
100001255	Benign	Female	African American	Status post hysterectomy
100001259	Benign	Female	White	Pre-menopausal
100001270	Benign	Female	White	Post-menopausal
100001279	Benign			
100001337	Benign	Female	White	Post-menopausal
100001352	Benign	Female	African American	Pre-menopausal
100001360	Benign	Female	White	Pre-menopausal
100001362	Benign	Female	White	Pre-menopausal
100001378	Benign	Female	African American	Pre-menopausal
100001392	Benign	Female	White	Pre-menopausal
100001393	Benign	Female	African American	Pre-menopausal
100001400	Benign	Female	Hispanic (white)	Pre-menopausal
100001407	Benign	Female	White	Pre-menopausal
100001408	Benign	Female	White	Post-menopausal
100001411	Benign	Female	African American	Pre-menopausal
100001413	Benign	Female	African American	Pre-menopausal
100001415	Benign	Female	Hispanic (white)	Pre-menopausal
100001416	Benign	Female	Other	Pre-menopausal
100001419	Benign	Female	White	Pre-menopausal
100001422	Benign	Male	White	
100001498	Benign	Female	African American	Pre-menopausal
100001507	Benign	Female	African American	Pre-menopausal
100001508	Benign	Female	White	Pre-menopausal
100001511	Benign	Female	African American	Pre-menopausal
100001624	Benign	Female	White	Pre-menopausal
100001630	Benign	Female	African American	Post-menopausal
100001634	Benign	Female	African American	Pre-menopausal
100001637	Benign	Female	African American	Pre-menopausal
100001648	Benign	Female	African American	Post-menopausal
100001659	Benign	Female	Hispanic (other)	Pre-menopausal
100001677	Benign	Female	White	Status post hysterectomy
100001687	Benign	Female	White	Pre-menopausal
100001693	Benign	Female	African American	Pre-menopausal
100001695	Benign	Male	White	
100001702	Benign	Female	White	Pre-menopausal

100001811	Benign	Female	African American	Pre-menopausal
100001827	Benign	Female	White	Pre-menopausal
100001848	Benign	Male	White	
100001883	Benign	Female	African American	Post-menopausal
100001888	Benign	Female	African American	Pre-menopausal
100001899	Benign	Female	White	Pre-menopausal
100001901	Benign	Female	African American	Status post hysterectomy
100001904	Benign			
100001905	Benign	Female	White	Pre-menopausal
100001977	Benign	Female	Hispanic (white)	Post-menopausal
100001995	Benign			
226	In situ	Female	African American	Post-menopausal
100000789	In situ	Female	White	Surgically menopausal
100001160	In situ			
100001986	In situ	Female	White	Surgically menopausal
187	Invasive	Female	Hispanic (other)	Post-menopausal
369	Invasive	Female	White	Post-menopausal
633	Invasive	Female	White	Post-menopausal
707	Invasive	Female	White	Post-menopausal
793	Invasive	Female	Hispanic (other)	Pre-menopausal
1197	Invasive	Female	White	Status post hysterectomy
200027	Invasive			
100001092	Invasive			
100001258	Invasive	Female	White	Pre-menopausal
100001354	Invasive	Female	White	Post-menopausal
100001361	Invasive	Female	White	Surgically menopausal
100001420	Invasive	Female	White	Post-menopausal
100001505	Invasive	Female	White	Post-menopausal
100001708	Invasive	Female	White	Post-menopausal
100001897	Invasive	Female	White	Post-menopausal
386	Malignant NOS	Female	White	Post-menopausal

We also performed gene expression study on breast tissues available from the above 92 subjects. 84 of them have breast tissues, with different types of diagnosis as shown below. To enable meaningful analysis, we only selected benign and invasive samples.

Sample	**Count**
Atypical 1	
Benign 53	
In Situ	4
Invasive 26	
	84

Out of the 26 Invasive samples, 4 were unusable (tumor exhausted or sample is not available). While processing samples, we found that 10 of these samples have tumor exhausted. Finally, only 12 samples were available for the project and these were used for

the gene expression analysis**.** From the 53 benign samples, we selected only 25 samples for the project with the following criteria,

Patient diagnosis should be – Benign
Patient status when samples were drawn should be – Benign
Sample diagnosis also should be – Benign.

Overall, for tissue samples we have gene expression data for 37 samples.

CBCP-ID	TYPE
100000912	Benign
100001977	Benign
100001905	Benign
100001888	Benign
100001862	Benign
100001848	Benign
100001702	Benign
100001693	Benign
100001687	Benign
100001637	Benign
100001634	Benign
100001630	Benign
100001624	Benign
100001422	Benign
100001416	Benign
100001415	Benign
100001408	Benign
100001407	Benign
100001393	Benign
100001253	Benign
100001101	Benign
100001090	Benign
100001087	Benign
100001047	Benign
528	Benign
100001897	Invasive
100001841	Invasive
100001420	Invasive
100001364	Invasive
100001361	Invasive
100001064	Invasive
707	Invasive
633	Invasive
416	Invasive
344	Invasive
278	Invasive
123	Invasive

II.3. Gene expression experiments

II.3.1 Task Description

Affymetrix GeneChip Human Genome U133 Plus 2.0 Array (Affymetrix, Santa Clara, CA, USA) were used for this study. It is a single array with over 47,000 transcript probe sets representing over 38,500 well-substantiated Human genes. The Affymetrix Human Genome U133-2.0 plus GeneChip, contains more than 54,000 probe sets representing greater than 47,000 transcripts, derived from approximately 38,500 well-substantiated human genes.

For blood, total RNA is extracted from whole blood using the PAXgene™ Blood RNA System Kit employing the manufacturer's instructions (PreAnalytiX) . RNA quality is determined with the Agilent 2100 Bioanalyzer (Agilent Technologies). The total RNA is cleared of the globin mRNA using the Globinclear kit (Ambion). The sample is ready for microarray analysis described next to the subsequent paragraph.

For invasive tissues, a single tumor sample is assayed. Laser capture microdissection is used to separate tumor tissue from surrounding stroma. RNA is isolated from microdissected samples (RNAqueous-micro kit). The isolated RNA is amplified, labeled, fragmented and hybridized to the microarray chip as described above.

For benign tissues, RNA is extracted from tissue sections using the RNeasy Mini Kit (Qiagen). The isolated RNA is amplified, labeled, fragmented and hybridized to the microarray chip as described below. Note that the amplification is performed in two rounds, starting from 10ng for the first and then 1 ug for the second.

We followed the Affymetrix GeneChip® Expression Analysis Technical Manual for all GeneChip array procedures. Briefly, one μg of total RNA is used for reverse transcription to produce single strand cDNA followed by second strand synthesis to form double strand cDNA. After cDNA purification, biotin-labeled aRNA target is produced by an *in vitro* transcription (IVT) reaction using the cDNA template. After aRNA purification, an aliquot of the labeled aRNA is run on Agilent's Bioanalyzer as a quality check and another aliquot is quantified using the Nanodrop UV/Vis spectrophotometer (Nanodrop). Only high quality aRNA with a yield of more than 15 μg is fragmented and hybridized to Affymetrix GeneChip arrays overnight (16 hours) in a temperature-controlled hyb-oven. After hybridization, GeneChip arrays are loaded onto a Fluidic Station 450 for washing and staining using the standard Affymetrix procedure. After the final wash, the GeneChip arrays are scanned using the Affymetrix GeneChip scanner 3000 G7. Scanned images are analyzed using Affymetrix data analysis software (GDAS) to generate the raw data.

II.3.2Data analysis

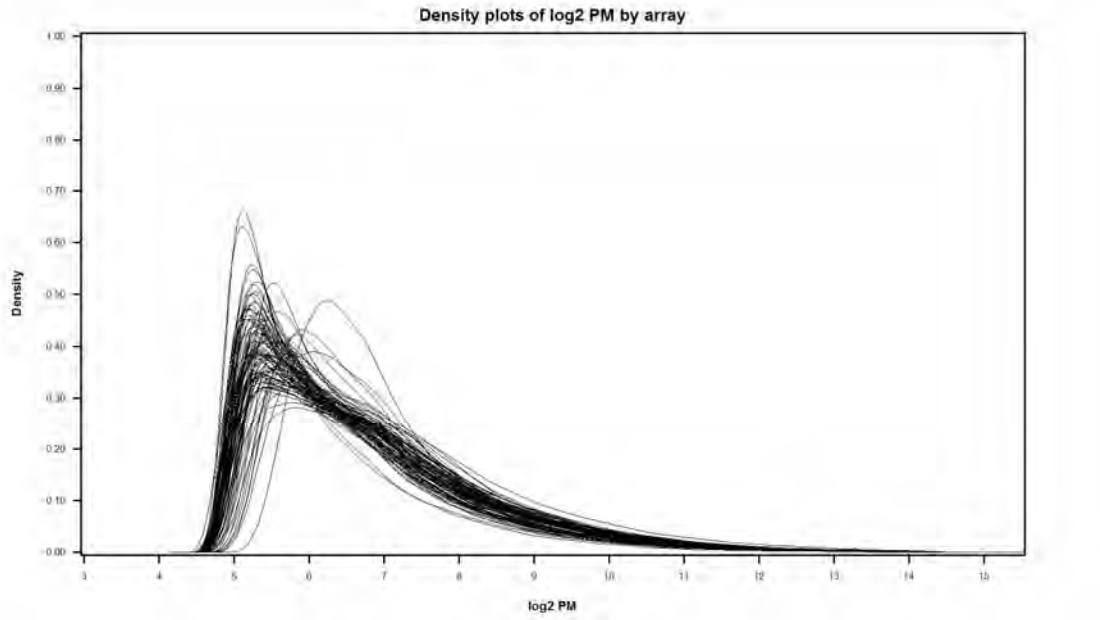
Data analysis is done usingR, SAS and GenomatixSuite PE and other statistical and Bioinformatics techniques and software.

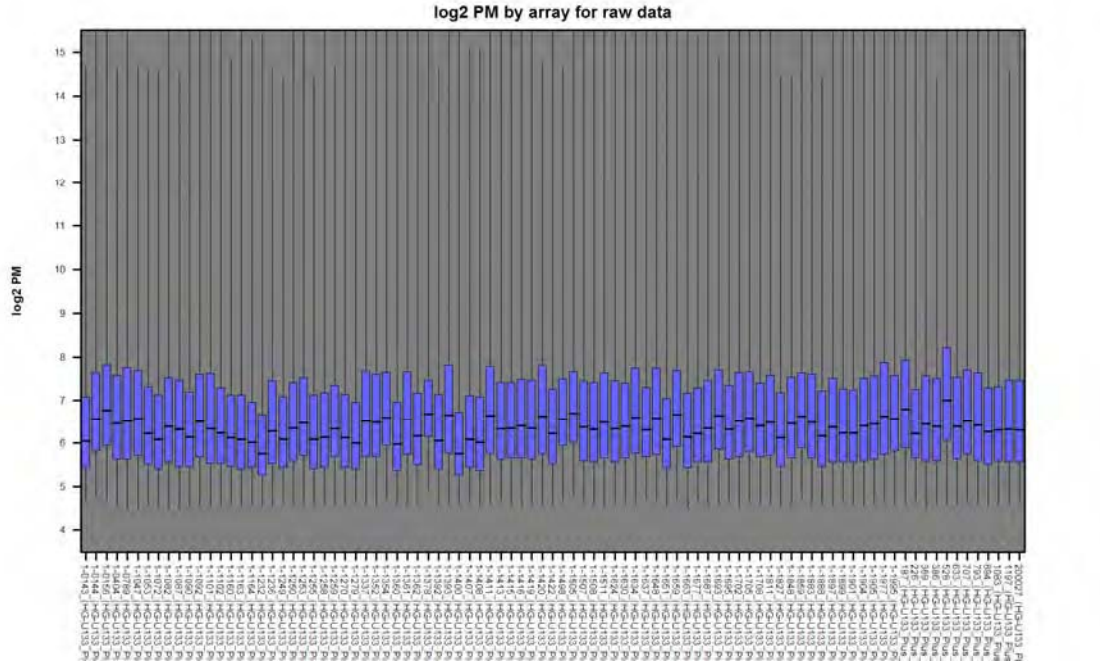
II.5. Results

The data analysis of the gene expression microarray experiments is divided into two section, with the first section focusing on blood sample data analysis and the second focusing on tissue sample analysis.

II.3.3 Blood sample gene expression data analysis

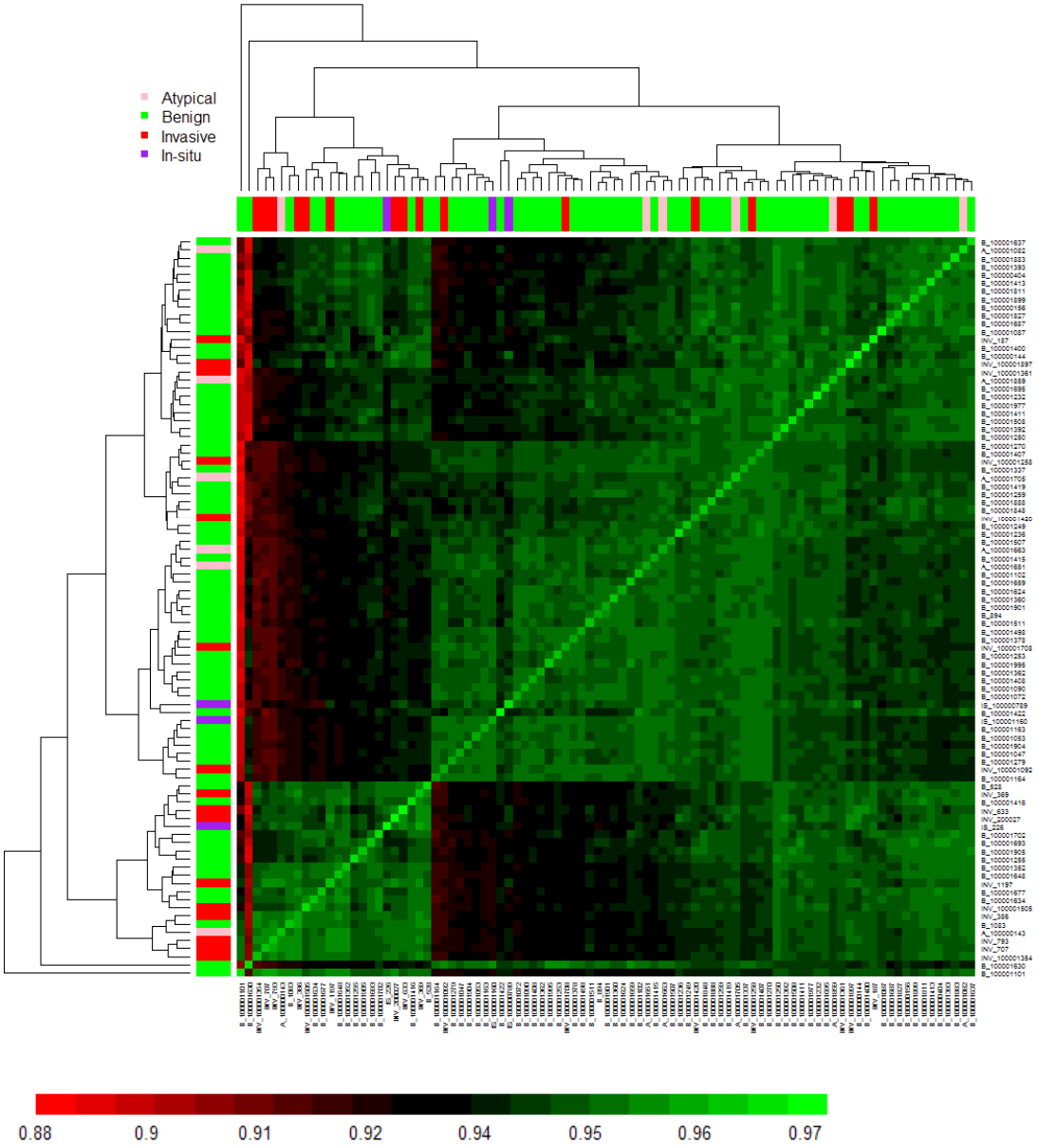
We first plot the raw data and check the distributions.
Raw data visualization:





From both the density plots and box plots for the raw data, we can see there is considerate variation between chips in this study. Before any comparisons were carried out between these chips, we need to normalize the data across all chips in this dataset and make them comparable.
The method we use for the background correction, normalization and summarization for microarray data is Robust Multichip Average methods (RMA). This method does the background correction using probe level model, quantile normalization, and probeset expression level summarization using median polish method.
After normalization, residual images were plot for each chip for quality assessment. The chips with apparent artifact were removed from the following analysis. To summarize, in this microarray data set using blood samples, we have 91 chips for the following analysis.

Visualization of correlation between the different samples:
One of the important issues in microarray data analysis is to investigate the similarity or the correlation between the samples. One way to do this is to calculate the correlation matrix and visualize it to see the correlation between the samples. In our dataset, we calculate the Pearson correlation matrix and visualize the results in a heat map. See the figure below.

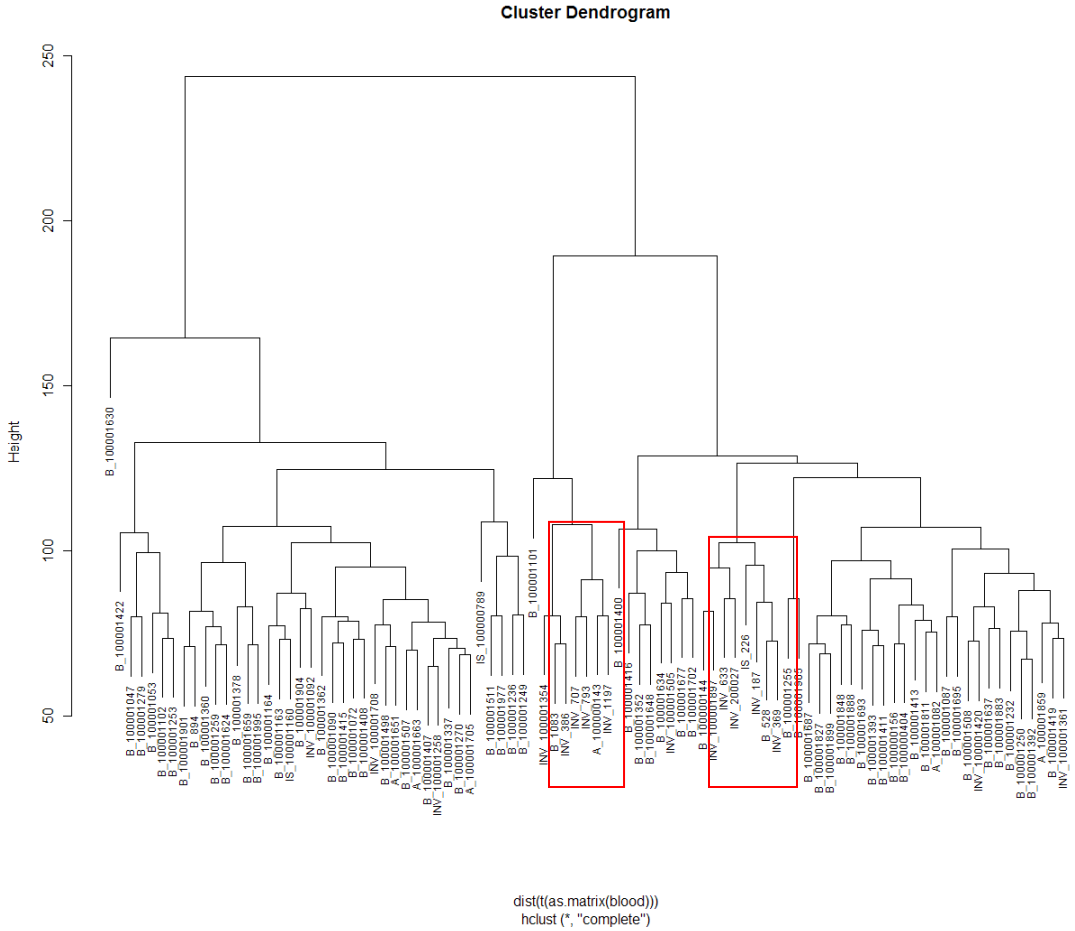


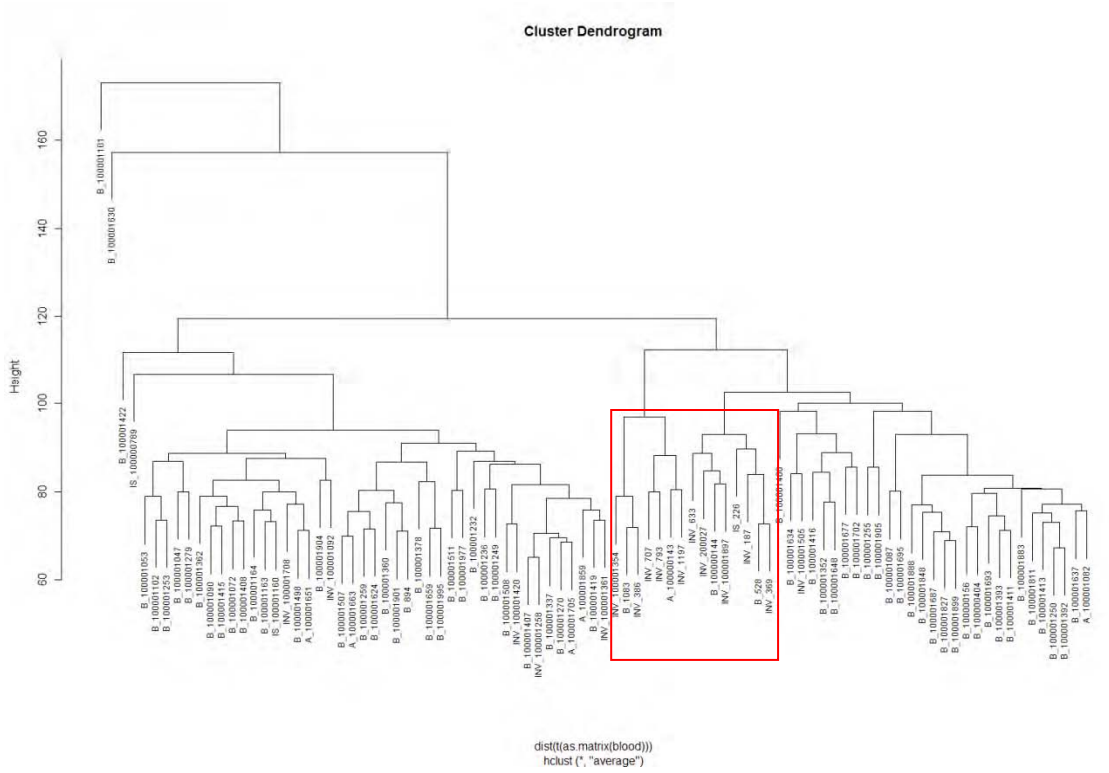
In the correlation heat map, the rows and column show the samples. In the row side and column side color bars represent the samples in different categories. With pink show the "Atypical" group, green is "Benign", red is "Invasive" , and purples are "In-situ". The bottom color key represents the correlation in the samples, with red for low correlation and green for higher correlation.

From the results of the correlation matrix, there are three major groups. Most of the invasive sample clustered in the lower-left square (9/16 invasive sample in this cluster). This cluster also includes one atypical sample and one "in-situ" samples. The other 7 invasive samples are randomly distributed in the benign groups. Most of atypical (pink) samples are clustered together with benign samples (5/6). Furthermore, for three in-situ samples, two are cluster with benign groups.

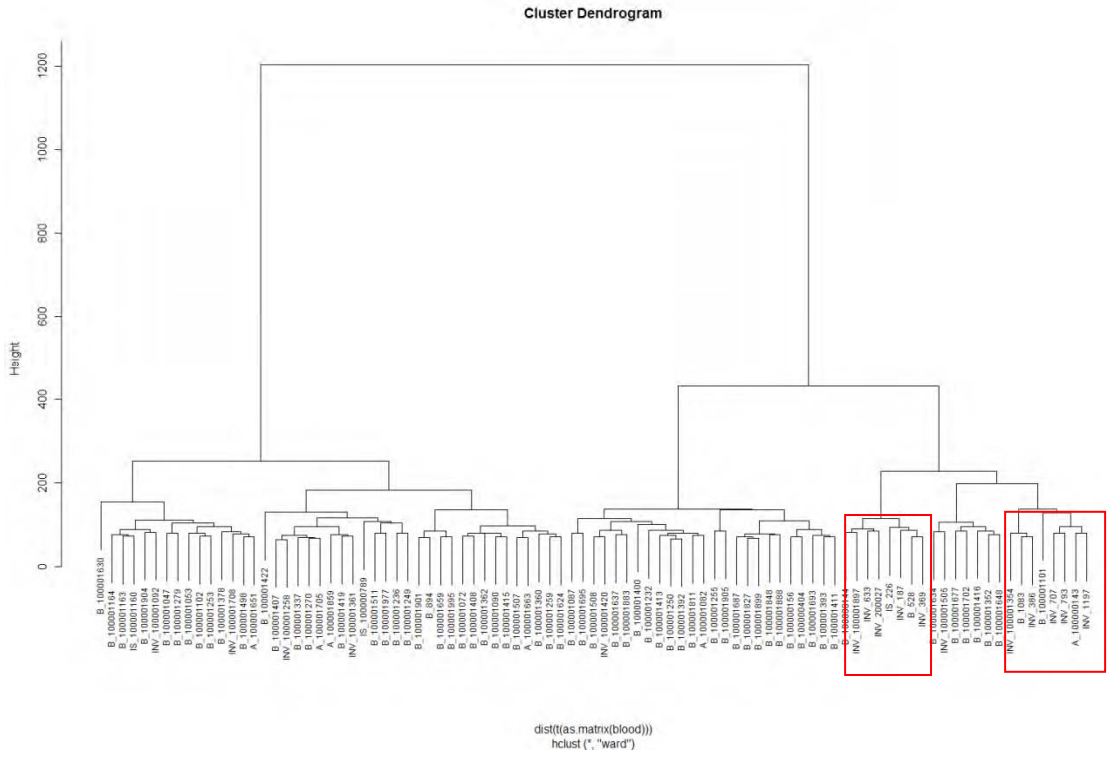
<u>Hierarchical clustering analysis on gene expression profiles:</u>

Clustering methods are widely applied in the microarray gene expression study. In this data analysis, we applied different clustering algorithms for our microarray data with blood samples. Four results were visualized by dendrogram and shown here using four hierarchical clustering methods (the algorithms are “complete”, “average”, “ward” and “mcquitty”). All these dendrograms show the consistent results, in which the majority of invasive samples (10/16) are clustered into two subclusters (which are highlighted by red rectangles).

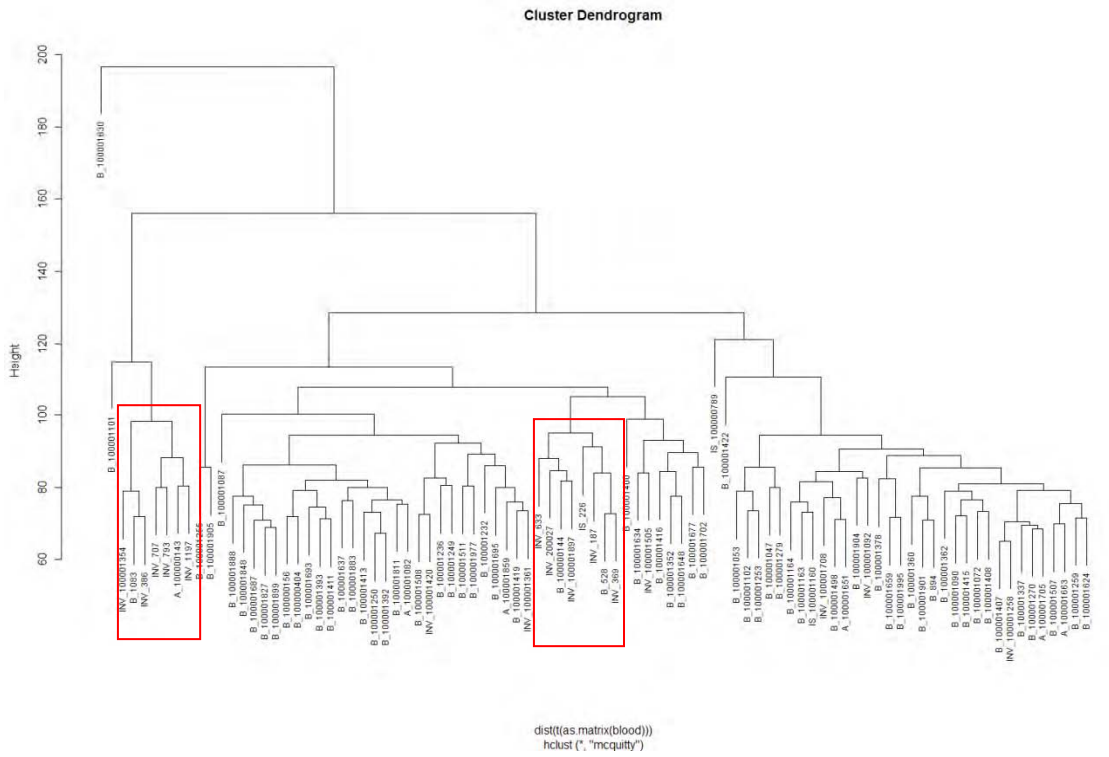




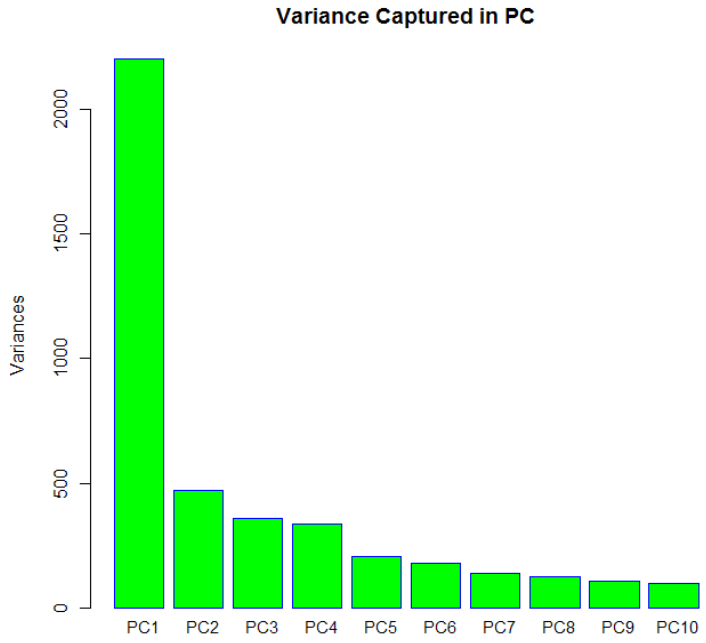
Algorithm: ward

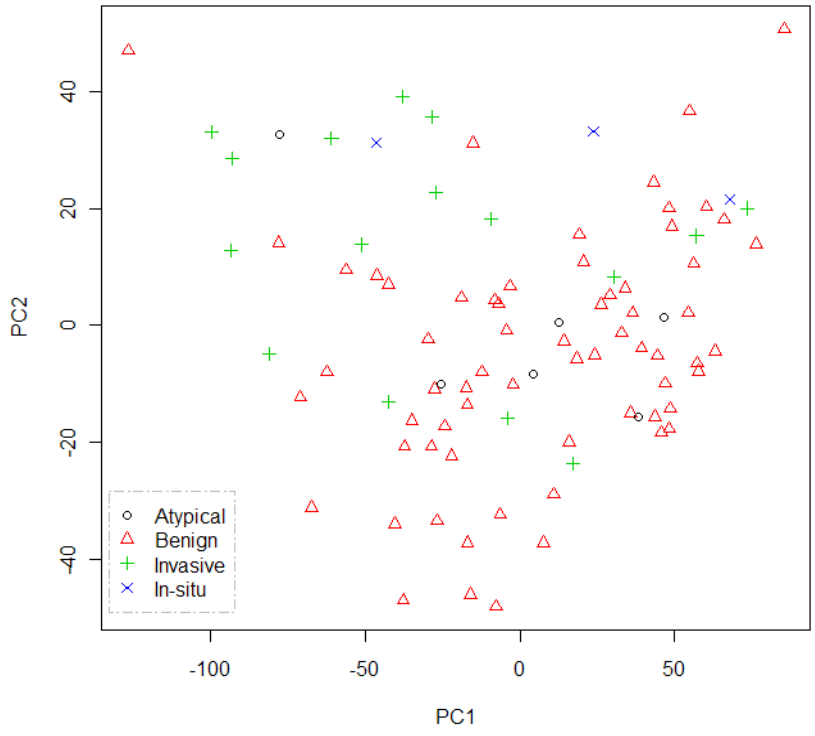


Algorithm: mcquitty



PCA for sample classification:
Microarray data is highly multi-dimensional with fewer sample numbers (from tens to hundreds) and tens of thousands of genes (variables). Principal Component Analysis (PCA) is a well-known method for displaying the pattern in the data by reducing the dimensions, which captures the variations in the first few components by linear combination.
We show the variation graph and also plot PC1 vs PC2 (see the following graphs). Both graphs show that the majority of variation in this data set. The PC1 vs PC2 plot shows that the samples from invasive and benign group are relatively separated well. Furthermore, most of the Atypical samples (except one sample) were classified with benign groups; and In-situ samples are similar to invasive groups.





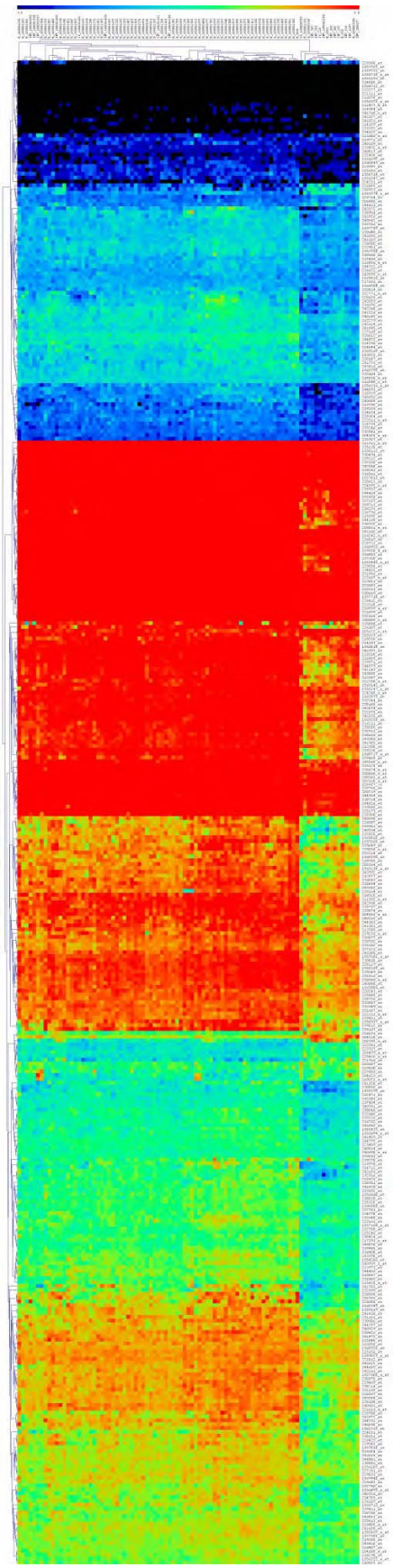
Differentially expression genes and patterns:

We have preformed statistical analysis to identify the differentially expressed genes. We summarized some analyses (based on different p-values or FDR control) and show some heatmaps below:
Analysis A:
Test: Kruskal-Wallis Test
Significance Based on Est. FDR (Benjamini-Hochberg)
Selected FDR Limit: 0.2
Computed FDR for Sig. Genes: 0.19976817

Group Information:
Atypical (6 samples in analysis)
Benign (66 samples in analysis)
In-situ (3 samples in analysis)
Invasive (16 samples in analysis)

Significant genes # of Significant Genes: 392
% of Genes that are Signficant: 1%

Non-significant genes # of non-significant Genes: 54283
% of Genes that are not signficant: 99%
Note: FDR control is used Benjamini-Hochberg method.



Analysis B:
Test: Kruskal-Wallis Test
Significance Based on Input Alpha
alpha: $p < 0.0050$

Group Information:
Atypical (6 samples in analysis)
Benign (66 samples in analysis)
In-situ (3 samples in analysis)
Invasive (16 samples in analysis)

Significant genes # of Significant Genes: 999
% of Genes that are Signficant: 2%

Non-significant genes # of non-significant Genes: 53676
% of Genes that are not signficant: 98%

Analysis C:
Test: Kruskal-Wallis Test
Significance Based on Input Alpha
alpha: $p < 0.0010$

Group Information:
Atypical (6 samples in analysis)
Benign (66 samples in analysis)
In-situ (3 samples in analysis)
Invasive (16 samples in analysis)

Significant genes # of Significant Genes: 290
% of Genes that are Signficant: 1%

Non-significant genes # of non-significant Genes: 54385
% of Genes that are not signficant: 99%

Final:
Benign vs Invasive (blood samples) only:
Test: Wilcoxon Rank Sum Test
Significance Based on Input Alpha
alpha: $p < 0.0010$

Group Information:
Benign (66 samples in analysis)
Invasive (16 samples in analysis)

Significant genes # of Significant Genes: 961

% of Genes that are Signficant: 2%

Non-significant genes # of non-significant Genes: 53714

% of Genes that are not signficant: 98%

We show the genes with FC greater than 1.5 in the following table:

Probesets GENE	TITLE	GENE_SYMBOL	ENTREZ_ID	UNIGENE_ID p	-value	AVE_B	AVE_I FC	
205849_s_at	ubiquinol-cytochrome c reductase binding protein	UQCRB	7381 Hs	.131255	3.75E-04 6.	84	8.69	-3.6
205681_at	BCL2-related protein A1	BCL2A1	597	Hs.227817	7.85E-04 7.	19	8.80	-3.1
217256_x_at	hCG1789827 | similar to large subunit ribosomal protein L36a	LOC641903 | LOC643505 | LOC646175 | LOC649299 | LOC732102 | hCG_1789827	641903 | 643505 | 646175 | 649299 | 728202 | 732102 Hs	.693269	8.90E-04 6.	81	8.42	-3.0
225312_at	COMM domain containing 6	COMMD6	170622	Hs.508266	7.85E-04 7.	55	9.07	-2.9
214512_s_at	SUB1 homolog (S. cerevisiae)	SUB1	10923	Hs.229641	5.11E-04 7.	27	8.77	-2.8
201304_at	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 5, 13kDa	NDUFA5	4698	Hs.651219	9.28E-04 4.	31	5.76	-2.7
1552701_a_at	caspase-1 dominant-negative inhibitor pseudo-ICE	COP1	114769	Hs.348365	6.62E-04 7.	02	8.43	-2.6
202635_s_at	polymerase (RNA) II (DNA directed) polypeptide K, 7.0kDa	POLR2K	5440	Hs.351475	4.89E-04 5.	31	6.64	-2.5
222465_at	chromosome 15 open reading frame 15 | similar to ribosomal protein L24-like	C15orf15 | LOC284288	284288 | 51187 H	s.274772	7.85E-04 5.	87	7.19	-2.5
212537_x_at	chromosome 18 open reading frame 32 | hCG22804 | hCG39912 | ribosomal protein L17	C18orf32 | RPL17 | hCG_22804 | hCG_39912	497661 | 6139 | 642250 | 645441 Hs	.293653	7.52E-04 9.	52	10.82	-2.5
212270_x_at	chromosome 18 open reading frame 32 | hCG22804 | hCG39912 | ribosomal protein L17	C18orf32 | RPL17 | hCG_22804 | hCG_39912	497661 | 6139 | 642250 | 645441 Hs	.293653	9.28E-04 9.	45	10.74	-2.4
205041_s_at	orosomucoid 1 | orosomucoid 2	ORM1 | ORM2	5004 | 5005	Hs.567311	5.57E-04 5.	65	6.93	-2.4
209795_at CD	69 molecule	CD69	969	Hs.208854	6.34E-04 4.	27	5.52	-2.4
224587_at	SUB1 homolog (S. cerevisiae)	SUB1	10923	Hs.229641	6.62E-04	5.94 7.	11	-2.3
223480_s_at	mitochondrial ribosomal protein L47 MRPL	47	57129	Hs.283734	3.28E-04 6.	40	7.57	-2.2
225207_at	pyruvate dehydrogenase kinase, isozyme 4	PDK4	5166	Hs.8364	7.21E-04 5.	74	6.82	-2.1
217147_s_at	T cell receptor associated transmembrane adaptor 1	TRAT1 50	852	Hs.138701	6.62E-04	6.16 7.	24	-2.1
200717_x_at	ribosomal protein L7	RPL7	6129	Hs.571841	8.19E-04 10.	98	12.06	-2.1
205040_at oro	somucoid 1	ORM1	5004	Hs.567311	5.11E-04 6.	33	7.40	-2.1
212042_x_at	hCG31916 | ribosomal protein L7	RPL7 | hCG_31916	6129 | 653702 Hs	.421257	6.91E-04 10.	51	11.57	-2.1
227840_at	hypothetical protein LOC130355	LOC130355	130355	Hs.99488	2.28E-04 3.	82	4.87	-2.1
209303_at	NADH dehydrogenase (ubiquinone) Fe-S protein 4, 18kDa (NADH-coenzyme Q reductase)	NDUFS4 4	724	Hs.528222	7.52E-04 6.	24	7.29	-2.1
218830_at	ribosomal protein L26-like 1	RPL26L1	51121	Hs.546390	5.57E-04 6.	58	7.59	-2.0
208808_s_at	high-mobility group box 2	HMGB2	3148	Hs.434953	7.85E-04 8.	35	9.36	-2.0
200099_s_at	ribosomal protein S3A | similar to ribosomal protein S3a	LOC439992 | RPS3A	439992 | 6189 Hs	.356572	4.68E-04 10.	97	11.97	-2.0
241705_at	ATP-binding cassette, sub-family A (ABC1), member 5	ABCA5	23461	Hs.421474	1.18E-04 6.	44	5.44	2.0
208995_s_at	peptidylprolyl isomerase G (cyclophilin G)	PPIG	9360	Hs.470544	3.81E-05 5.	59	6.58	-2.0
201012_at ann	exin A1	ANXA1	301	Hs.494173	3.28E-04 9.	43	10.40	-2.0

202917_s_at	S100 calcium binding protein A8 S1	00A8	6279	Hs.416073	3.59E-04	11.92	12.88	-1.9
201595_s_at	zinc finger CCCH-type containing 15	ZC3H15	55854	Hs.696083	7.21E-04	6.28	7.23	-1.9
217491_x_at	cytochrome c oxidase subunit VIIc COX7	C	1350	Hs.430075	8.90E-04	8.71	9.65	-1.9
203543_s_at	Kruppel-like factor 9	KLF9	687	Hs.150557	1.81E-04	3.16	4.10	-1.9
1552772_at	C-type lectin domain family 4, member D	CLEC4D	338339	Hs.351811	4.89E-04	5.68	6.61	-1.9
214709_s_at	kinectin 1 (kinesin receptor)	KTN1	3895	Hs.509414	4.48E-04	6.84	7.76	-1.9
224914_s_at	DnaJ (Hsp40) homolog, subfamily C, member 14 | cytokine induced protein 29 kDa	CIP29 | DNAJC14	84324 | 85406	Hs.505676	6.07E-04	5.55	6.45	-1.9
240594_at	Transcribed locus	NA	NA	Hs.668170	1.81E-04	7.08	6.20	1.8
200915_x_at	kinectin 1 (kinesin receptor)	KTN1	3895	Hs.509414	4.89E-04	7.12	8.00	-1.8
232882_at	CDNA FLJ12289 fis, clone MAMMA1001788 NA		NA	Hs.687769	2.28E-04	8.59	7.72	1.8
200093_s_at	histidine triad nucleotide binding protein 1	HINT1	3094	Hs.483305	8.54E-04	8.40	9.26	-1.8
219598_s_at	RWD domain containing 1 | hypothetical protein LOC727789	LOC727789 | RWDD1	51389 | 727789 Hs	.532164	2.08E-04	6.79	7.65	-1.8
201257_x_at	ribosomal protein S3A	RPS3A	6189	Hs.356572	9.68E-04	11.45	12.31	-1.8
217092_x_at	similar to 60S ribosomal protein L7 LOC6	46912	646912	Hs.648250	7.52E-04	7.19	8.04	-1.8
230923_at	family with sequence similarity 19 (chemokine (C-C motif)-like), member A1	FAM19A1 40	7738	Hs.655061	2.39E-04	4.09	4.94	-1.8
1565887_at	Transient receptor potential cation channel, subfamily M, member 7	TRPM7 54	822	Hs.512894	3.00E-04	6.01	5.17	1.8
212391_x_at	ribosomal protein S3A	RPS3A	6189	Hs.356572	7.21E-04	11.49	12.34	-1.8
225580_at	mitochondrial ribosomal protein L50 MRPL	50	54534	Hs.288224	1.30E-04	2.71	3.56	-1.8
244414_at	NA	NA	NA	NA	2.55E-06	9.54	8.71	1.8
1556932_at	Full length insert cDNA YH97G12	NA NA		Hs.633173	2.62E-04	6.85	6.02	1.8
232307_at	CDNA FLJ11492 fis, clone HEMBA1001939 NA		NA	Hs.656085	7.52E-04	9.01	8.18	1.8
201139_s_at	Sjogren syndrome antigen B (autoantigen La)	SSB	6741	Hs.632535	5.11E-04	6.15	6.98	-1.8
236307_at	Transcribed locus	NA	NA	Hs.660736	1.08E-04	8.07	7.26	1.8
1560342_at	CDNA clone IMAGE:5275043	NA	NA	Hs.684396	7.02E-06	6.76	5.95	1.8
242405_at	Transcribed locus	NA	NA	Hs.662061	3.00E-04	8.93	8.12	1.7
239798_at	Transcribed locus	NA	NA	Hs.660359	6.07E-04	7.18	6.37	1.7
201795_at	lamin B receptor	LBR	3930	Hs.435166	7.21E-04	9.26	10.06	-1.7
221805_at	neurofilament, light polypeptide 68kDa N	EFL	4747	Hs.521461	1.08E-04	3.64	4.44	-1.7
210538_s_at	baculoviral IAP repeat-containing 3	BIRC3	330	Hs.127799	7.52E-04	6.66	7.42	-1.7
242737_at	Transcribed locus	NA	NA	Hs.674001	9.28E-04	6.43	5.67	1.7
234330_at	CDNA FLJ14081 fis, clone HEMBB1002280	NA NA		Hs.573373	2.62E-04	5.20	4.45	1.7
241271_at	NA	NA	NA	NA	5.57E-04	4.92	4.17	1.7
201757_at	NADH dehydrogenase (ubiquinone) Fe-S protein 5, 15kDa (NADH-coenzyme Q reductase)	NDUFS5 4	725	Hs.632385	7.85E-04	8.83	9.58	-1.7
222156_x_at	cell cycle progression 1	CCPG1	9236	Hs.612814	6.07E-04	5.52	6.27	-1.7
221916_at	neurofilament, light polypeptide 68kDa N	EFL	4747	Hs.521461	1.57E-04	5.06	5.80	-1.7
232916_at	CDNA FLJ12097 fis, clone HEMBB1002617 NA		NA Hs.656531		2.40E-05	5.91	5.17	1.7

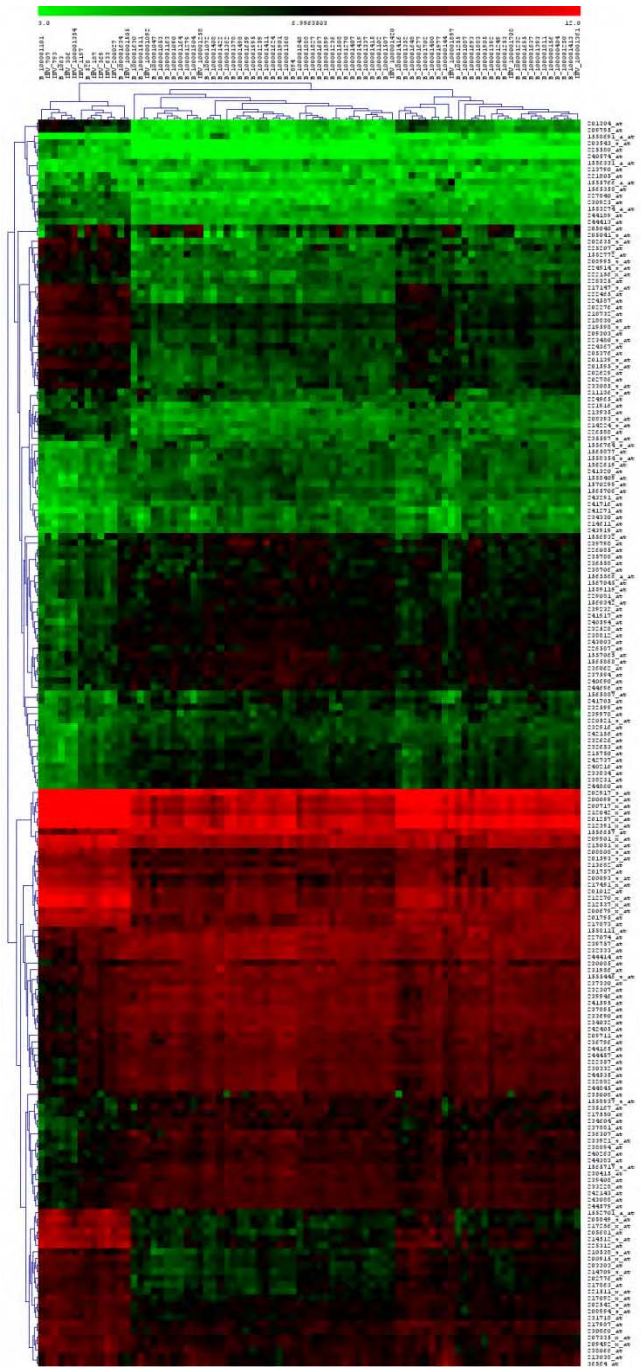
241320_at NA		NA	NA	NA	7.66E-05 5.	59	4.85	1.7
203303_at	dynein, light chain, Tctex-type 3	DYNLT3	6990	Hs.446392	9.28E-04 6.	88	7.61	-1.7
226588_at K	IAA1604 protein	KIAA1604	57703	Hs.311363	9.09E-04 5.	26	5.99	-1.7
233085_s_at	oligonucleotide/oligosaccharide-binding fold containing 2A	OBFC2A	64859	Hs.591610	9.68E-04 6.	32	7.04	-1.6
235167_at hy	pothetical gene LOC283846	DKFZp547E087	283846	Hs.648439	4.89E-04 7.	69	6.97	1.6
236558_at NA		NA	NA	NA	8.44E-05 6.	72	6.01	1.6
1556764_s_at	CDNA FLJ90128 fis, clone HEMBB1000276 NA		NA Hs.657837		2.28E-04 5.	52	4.80	1.6
202776_at	deoxynucleotidyltransferase, terminal, interacting protein 2	DNTTIP2	30836	Hs.85769	8.90E-04 6.	84	7.55	-1.6
232333_at	CDNA FLJ12284 fis, clone MAMMA1001757 NA		NA	Hs.658320	1.30E-04 9.	47	8.76	1.6
230085_at Tran	scribed locus	NA	NA	Hs.403937	2.18E-04	8.49 7.	79	1.6
244860_at Tran	scribed locus	NA	NA	Hs.610342	6.34E-04 6.	57	5.87	1.6
1555766_a_at	guanine nucleotide binding protein (G protein), gamma 2	GNG2	54331	Hs.695989	2.08E-04 4.	19	4.90	-1.6
238231_at	Nuclear transcription factor Y, gamma	NFYC 4	802	Hs.233458	2.18E-04	6.68 5.	98	1.6
224367_at	brain expressed X-linked 2	BEX2	84707	Hs.398989	9.68E-04 6.	39	7.09	-1.6
238066_at	retinol binding protein 7, cellular	RBP7	116362	Hs.422688	1.81E-04 7.	71	8.41	-1.6
217863_at	protein inhibitor of activated STAT, 1	PIAS1	8554	Hs.162458	8.54E-04 6.	92	7.61	-1.6
208994_s_at	peptidylprolyl isomerase G (cyclophilin G)	PPIG	9360	Hs.470544	2.86E-04 7.	31	8.00	-1.6
237594_at NA		NA	NA	NA	2.50E-04 7.	27	6.58	1.6
234604_at	CDNA: FLJ21228 fis, clone COL00739 NA		NA	Hs.677287	1.57E-04 7.	56	6.88	1.6
234032_at PRO1	550	NA	NA	Hs.684536	9.30E-05 9.	04	8.36	1.6
244845_at Tran	scribed locus	NA	NA	Hs.677811	5.82E-04 8.	59	7.91	1.6
237881_at Tran	scribed locus	NA	NA	Hs.653522	1.43E-04 7.	32	6.65	1.6
1556657_at	CDNA FLJ36459 fis, clone THYMU2014762 NA		NA	Hs.687293	5.33E-04	10.75 10.	08 1.	6
1557065_at	YLP motif containing 1	YLPM1	56252	Hs.531111	6.91E-04	7.32 6.	65	1.6
214224_s_at	protein (peptidylprolyl cis/trans isomerase) NIMA-interacting, 4 (parvulin)	PIN4 5	303	Hs.655623	8.90E-04	5.05 5.	73	-1.6
202276_at	split hand/foot malformation (ectrodactyly) type 1	SHFM1	7979	Hs.489201	6.07E-04 6.	60	7.26	-1.6
244189_at K	IAA1648 protein	KIAA1648	284900	Hs.602319	4.10E-04 4.	60	5.27	-1.6
237895_at NA		NA	NA	NA	4.10E-04 9.	19	8.52	1.6
209492_x_at	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit E	ATP5I 52	1	Hs.85539	3.59E-04	7.71 8.	38	-1.6
1570299_at	Homo sapiens, clone IMAGE:4702594, mRNA	NA	NA	Hs.681804	9.30E-05 5.	36	4.70	1.6
201593_s_at	zinc finger CCCH-type containing 15	ZC3H15	55854	Hs.696083	1.50E-04 8.	48	9.14	-1.6
230332_at	Zinc finger, CCHC domain containing 7	ZCCHC7	84186	Hs.654700	1.02E-04 8.	67	8.02	1.6
244535_at Tran	scribed locus	NA	NA	Hs.664595	6.62E-04 8.	59	7.93	1.6
221511_x_at	cell cycle progression 1	CCPG1	9236	Hs.612814	7.21E-04 7.	01	7.66	-1.6
218732_at	peptidyl-tRNA hydrolase 2	PTRH2	51651	Hs.12677	9.28E-04 6.	63	7.28	-1.6
208393_s_at	RAD50 homolog (S. cerevisiae)	RAD50	10111	Hs.655835	5.10E-04 4.	92	5.57	-1.6
238894_at Tran	scribed locus	NA	NA	Hs.659569	5.82E-04 7.	86	7.21	1.6
235597_s_at	RANBP2-like and GRIP domain containing 1 | RANBP2-like and	RGPD1 | RGPD2 | RGPD3	400966 | 653489 |	Hs.656849	4.89E-04 5.	27	5.92	-1.6

	GRIP domain containing 2 | RANBP2-like and GRIP domain containing 3		729857					
243291_at Tran	scribed locus	NA	NA	Hs.655845	7.85E-04	5.55	4.91	1.6
1558409_at	CDNA FLJ36478 fis, clone THYMU2017362 NA		NA	Hs.661830	4.28E-04	5.59	4.95	1.6
211136_s_at	cleft lip and palate associated transmembrane protein 1	CLPTM1	1209	Hs.444441	4.10E-04	5.90	6.54	-1.6
1565877_at	Full length insert cDNA clone YP86C01	NA NA		Hs.658484	6.07E-04	5.65	5.00	1.6
215750_at K	IAA1659 protein	KIAA1659	85373	Hs.675271	9.28E-04	6.40	5.76	1.6
1559119_at	CDNA FLJ25633 fis, clone STM04048 NA		NA	Hs.658775	1.02E-04	6.95	6.31	1.6
242156_at Tran	scribed locus	NA	NA	Hs.610345	8.73E-06	5.98	5.34	1.6
240690_at NA		NA	NA	NA	1.43E-04	7.25	6.61	1.6
232626_at	CDNA FLJ14143 fis, clone MAMMA1002892 NA		NA	Hs.657158	3.00E-04	6.23	5.59	1.6
239757_at	Zinc finger, AN1-type domain 6	ZFAND6	54469	Hs.654787	1.13E-04	9.59	8.95	1.6
233690_at	CDNA: FLJ23090 fis, clone LNG07119 NA		NA	Hs.677392	2.08E-04	9.01	8.37	1.6
222357_at	zinc finger and BTB domain containing 20	ZBTB20	26137	Hs.693802	3.92E-04	8.41	7.77	1.6
224965_at	guanine nucleotide binding protein (G protein), gamma 2	GNG2	54331	Hs.695989	7.52E-04	5.70	6.32	-1.5
243003_at	CDNA FLJ45369 fis, clone BRHIP3017325 NA		NA	Hs.657736	6.07E-04	7.09	6.46	1.5
209901_x_at	allograft inflammatory factor 1	AIF1	199	Hs.76364	9.68E-04	10.23	10.86	-1.5
200679_x_at	high-mobility group box 1	HMGB1	3146	Hs.644368	2.50E-04	9.35	9.98	-1.5
229081_at	Solute carrier family 25, member 13 (citrin)	SLC25A13	10165	Hs.489190	2.62E-04	6.88	6.26	1.5
238706_at	PAP associated domain containing 4	PAPD4 1	67153	Hs.418198	1.57E-04	6.81	6.18	1.5
1558937_s_at	MRNA (fetal brain cDNA b2_2g) NA		NA	Hs.677477	5.98E-05	7.56	6.94	1.5
235008_at	CDNA FLJ25241 fis, clone STM02689 NA		NA	Hs.658703	2.62E-04	7.94	7.31	1.5
202786_at	serine threonine kinase 39 (STE20/SPS1 homolog, yeast)	STK39	27347	Hs.276271	3.92E-04	6.13	6.76	-1.5
213682_at nu	cleoporin 50kDa	NUP50	10762	Hs.475103	4.28E-04	8.48	9.11	-1.5
207335_x_at	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit E | major facilitator superfamily domain containing 7	ATP5I | MFSD7	521 | 84179	Hs.567612	7.21E-04	7.54	8.16	-1.5
217987_at	asparagine synthetase domain containing 1	ASNSD1	54529	Hs.101364	6.34E-04	7.87	8.48	-1.5
242143_at Tran	scribed locus	NA	NA	Hs.611969	5.42E-05	8.02	7.41	1.5
231718_at	SLU7 splicing factor homolog (S. cerevisiae)	SLU7	10569	Hs.435342	7.52E-04	7.21	7.83	-1.5
1558354_s_at	CDNA clone IMAGE:5260583	NA	NA	Hs.672351	2.08E-04	5.57	4.95	1.5
232653_at	CDNA FLJ14044 fis, clone HEMBA1006124 NA		NA	Hs.688352	4.89E-04	6.30	5.69	1.5
240574_at	CDNA clone IMAGE:5262677	NA	NA	Hs.594844	5.11E-04	3.03	3.64	-1.5
1556331_a_at	CDNA clone IMAGE:5259142	NA	NA	Hs.658896	9.76E-05	4.42	3.82	1.5
244579_at Tran	scribed locus	NA	NA	Hs.673488	1.43E-04	8.11	7.51	1.5
1565717_s_at	fusion (involved in t(12;16) in malignant liposarcoma)	FUS	2521	Hs.513522	1.36E-04	8.10	7.50	1.5
238812_at Tran	scribed locus	NA	NA	Hs.306329	1.65E-04	7.03	6.43	1.5
220521_s_at	ATG16 autophagy related 16-like 1 (S. cerevisiae)	ATG16L1	55054	Hs.529322	9.68E-04	6.00	5.40	1.5
244165_at	chromosome 10 open reading frame 18	C10orf18	54906	Hs.432548	2.28E-05	8.59	7.99	1.5
241716_at	heat shock 60kDa protein 1 (chaperonin) H	SPD1	3329	NA	7.68E-04	5.16	4.56	1.5

209711_at	solute carrier family 35 (UDP-glucuronic acid/UDP-N-acetylgalactosamine dual transporter), member D1	SLC35D1	23169	Hs.213642	1.66E-05	8.71	8.11	1.5
202629_at	amyloid beta precursor protein (cytoplasmic tail) binding protein 2	APPBP2 105	13 H	s.84084	7.52E-04	6.05	6.65	-1.5
231956_at K	IAA1618	KIAA1618	57714	Hs.514554	2.62E-04	9.15	8.55	1.5
213935_at	abhydrolase domain containing 5 A	BHD5	51099	Hs.655670	7.85E-04	5.05	5.65	-1.5
226085_at	CDNA clone IMAGE:4842353	NA	NA	Hs.349283	2.86E-04	6.79	6.20	1.5
233834_at	CDNA: FLJ21392 fis, clone COL03505	NA NA		Hs.677315	2.62E-04	6.53	5.93	1.5
232599_at	exocyst complex component 6	EXOC6	54536	Hs.655657	1.65E-04	6.37	5.77	1.5
217550_at	activating transcription factor 6	ATF6	22926	Hs.492740	6.62E-04	7.54	6.95	1.5
237330_at Tran	scribed locus	NA	NA	Hs.663957	8.86E-05	8.93	8.34	1.5
239978_at NA		NA	NA	NA	6.61E-05	6.08	5.49	1.5
36564_at	ring finger protein 19B	RNF19B	127544	Hs.591504	2.18E-04	7.73	8.32	-1.5
243088_at NA		NA	NA	NA	6.61E-05	8.04	7.45	1.5
1567045_at	Full length insert cDNA clone YN86A01 NA		NA	Hs.658131	9.28E-04	7.04	6.46	1.5
230415_at	CDNA FLJ12381 fis, clone MAMMA1002566 NA		NA	Hs.656237	4.28E-04	8.12	7.53	1.5
1558111_at m	uscleblind-like (Drosophila)	MBNL1	4154	Hs.478000	2.86E-04	9.44	8.85	1.5
239408_at Tran	scribed locus	NA	NA	Hs.687626	7.21E-04	7.92	7.34	1.5
217873_at	calcium binding protein 39	CAB39	51719	Hs.632536	6.62E-04	9.05	9.62	-1.5
226587_at	CDNA FLJ33569 fis, clone BRAMY2010317 NA		NA	Hs.592473	4.28E-04	7.05	6.48	1.5
1565358_at	retinoic acid receptor, alpha	RARA	5914	Hs.654583	9.68E-04	3.99	4.56	-1.5
236062_at Tran	scribed locus	NA	NA	Hs.656820	5.94E-04	7.27	6.69	1.5
232528_at	CDNA FLJ11226 fis, clone PLACE1008280 NA		NA	Hs.661131	1.18E-04	7.09	6.52	1.5
1565566_a_at	Full length insert cDNA YN68A11 NA		NA	Hs.657994	6.62E-04	6.71	6.14	1.5
233921_s_at	CDNA FLJ12016 fis, clone HEMBB1001707 NA		NA Hs.671107		6.34E-04	7.83	7.25	1.5
244696_at NA		NA	NA	NA	9.68E-04	7.03	6.46	1.5
228325_at K	IAA0146	KIAA0146	23514	Hs.381058	3.75E-04	5.56	6.13	-1.5
236796_at	BTB and CNC homology 1, basic leucine zipper transcription factor 2	BACH2	60468	Hs.269764	1.36E-04	8.59	8.03	1.5
214611_at	glutamate receptor, ionotropic, kainate 1	GRIK1	2897	Hs.695938	7.84E-04	5.09	4.53	1.5
1565868_at	CD44 molecule (Indian blood group) CD	44	960	Hs.502328	4.22E-05	7.24	6.67	1.5
1562619_at	thioredoxin domain containing 6	TXNDC6	347736	Hs.660992	4.48E-04	5.54	4.98	1.5
1565706_at	CDNA: FLJ21395 fis, clone COL03557 NA		NA	Hs.677316	6.34E-04	5.68	5.12	1.5
240263_at Tran	scribed locus	NA	NA	Hs.687488	6.77E-05	7.66	7.09	1.5
239946_at Tran	scribed locus	NA	NA	Hs.687851	6.91E-04	8.87	8.31	1.5
205376_at	inositol polyphosphate-4-phosphatase, type II, 105kDa	INPP4B	8821	Hs.658245	3.43E-04	6.19	6.75	-1.5
244457_at Tran	scribed locus	NA	NA	Hs.677790	8.44E-05	8.53	7.97	1.5
227074_at	MRNA; cDNA DKFZp667D2123 (from clone DKFZp667D2123)	NA NA		Hs.648647	2.18E-04	9.61	9.05	1.5
241917_at Tran	scribed locus	NA	NA	Hs.673939	9.76E-05	6.82	6.26	1.5

233228_at	CDNA: FLJ21229 fis, clone COL00740 NA		NA	Hs.677288	4.01E-05	8.23	7.67	1.5
235788_at	Tran scribed locus	NA	NA	Hs.656029	5.42E-05	6.87	6.31	1.5
1553274_a_at	chromosome 6 open reading frame 151	C6orf151	154007	Hs.13366	3.83E-04	4.45	5.00	-1.5
202542_s_at	small inducible cytokine subfamily E, member 1 (endothelial monocyte-activating)	SCYE1 92	55	Hs.591680	8.72E-04	7.40	7.95	-1.5
213038_at	ring finger protein 19B	RNF19B	127544	Hs.591504	1.90E-04	7.61	8.16	-1.5
215051_x_at	allograft inflammatory factor 1	AIF1	199	Hs.76364	8.19E-04	10.17	10.72	-1.5
244413_at	C-t ype lectin-like 1	CLECL1	160365	Hs.560087	1.08E-04	4.20	4.75	-1.5
240216_at	Tran scribed locus	NA	NA	Hs.659543	3.59E-04	6.51	5.96	1.5
230860_at	Tran scribed locus	NA	NA	Hs.282800	4.48E-04	7.75	8.30	-1.5
244383_at	NA	NA	NA	NA	3.81E-05	7.65	7.10	1.5
239232_at	Musashi homolog 2 (Drosophila)	MSI2 1	24540	Hs.658922	5.57E-04	6.89	6.35	1.5
1558691_a_at	dedicator of cytokinesis 4	DOCK4	9732	Hs.654652	1.43E-04	3.39	2.84	1.5
243919_at	Tran scribed locus	NA	NA	Hs.687900	8.19E-04	5.15	4.60	1.5
1555446_s_at	transmembrane protein 1	TMEM1	7109	Hs.126221	5.11E-04	9.03	8.49	1.5
241595_at	NA	NA	NA	NA	6.34E-04	8.96	8.42	1.5
213790_at	ADAM metallopeptidase domain 12 (meltrin alpha)	ADAM12	8038	Hs.655388	1.75E-05	4.00	3.46	1.5

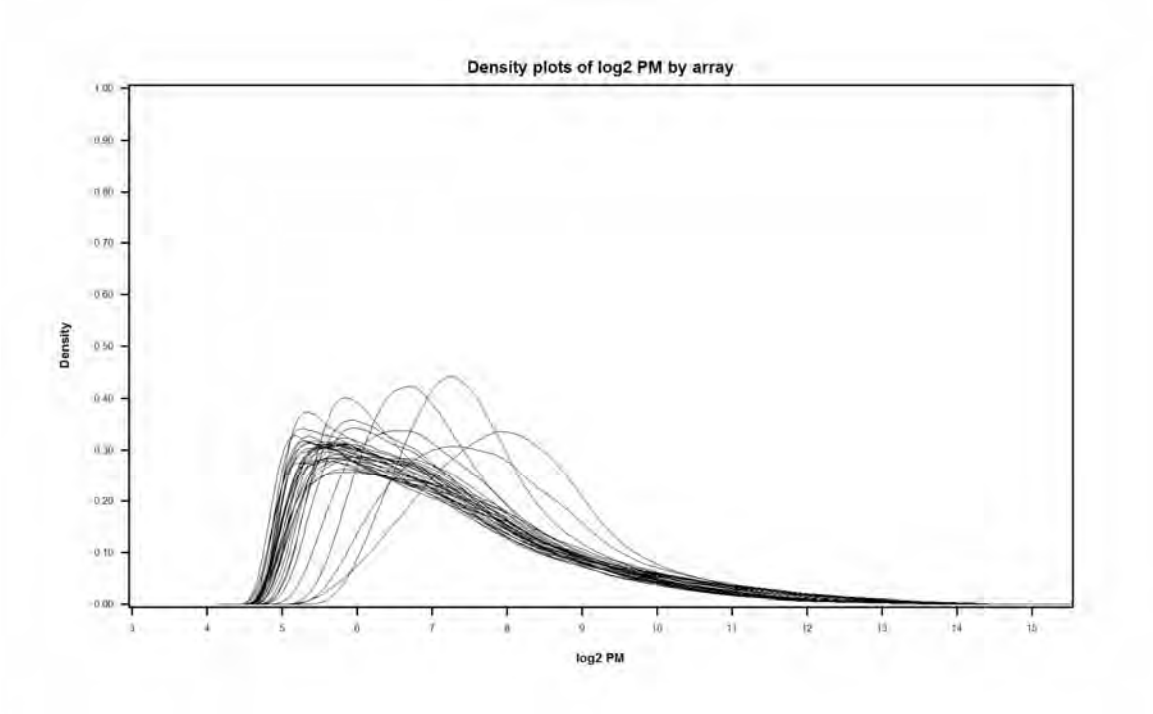
The following heatmap shows the clustering results for the these genes.



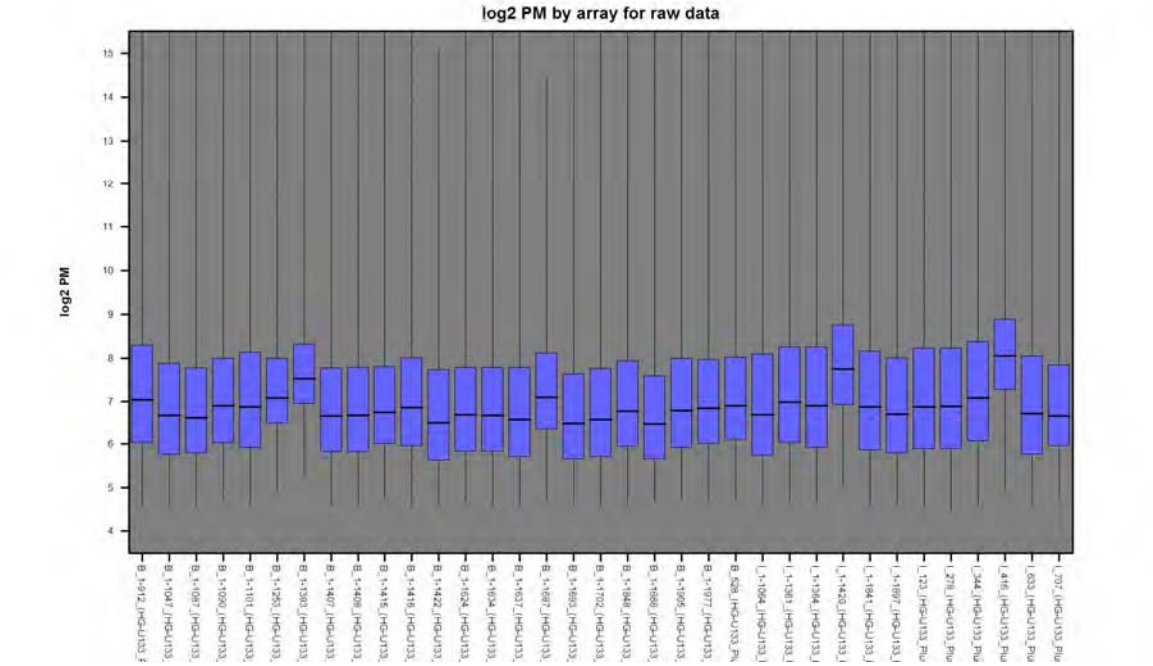
II.3.4 Tissue sample gene expression data analysis

Totally, 37 tissue samples are used for microarray experiment in this study; 25 of them are benign, and 12 are invasive (see the table below for CBCP-ID and patients categories: benign or invasive). Microarray analysis are preformed as described in the previous blood sample sections.

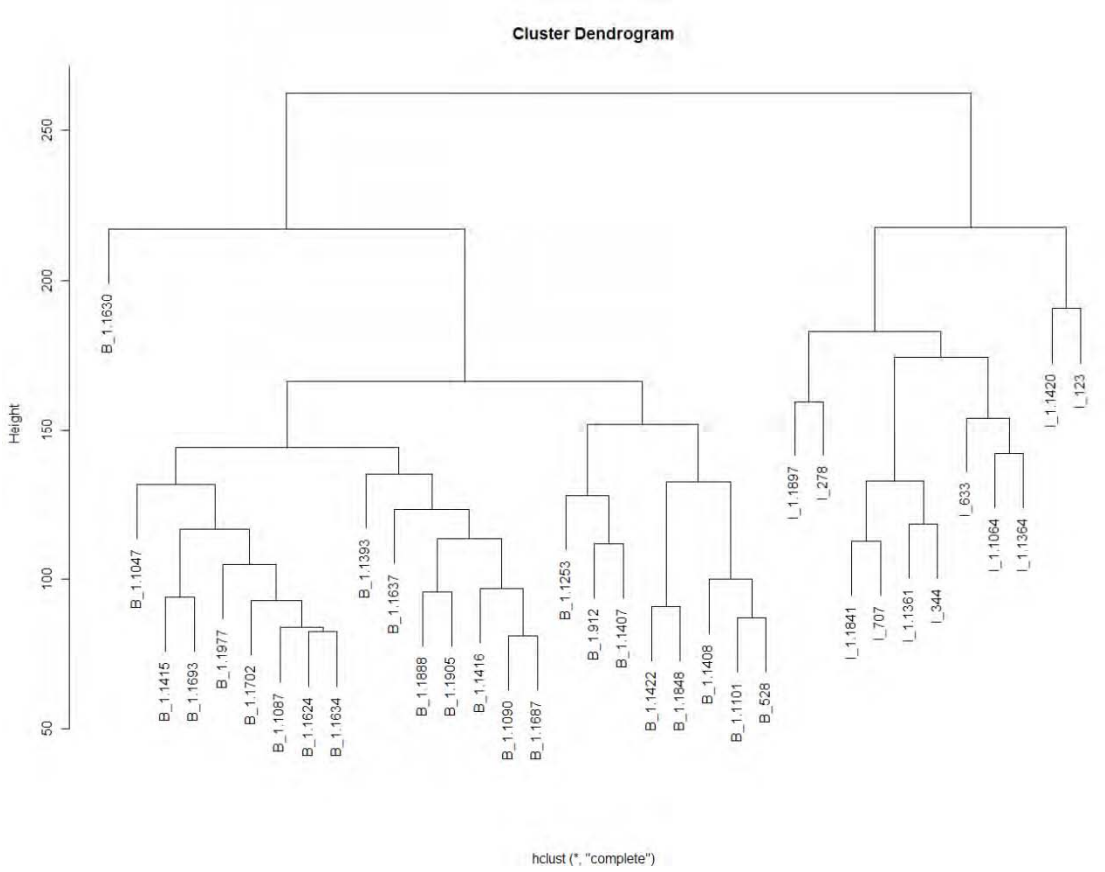
The analysis process is same as described above. Boxplot and density plot for gene expression profiling:



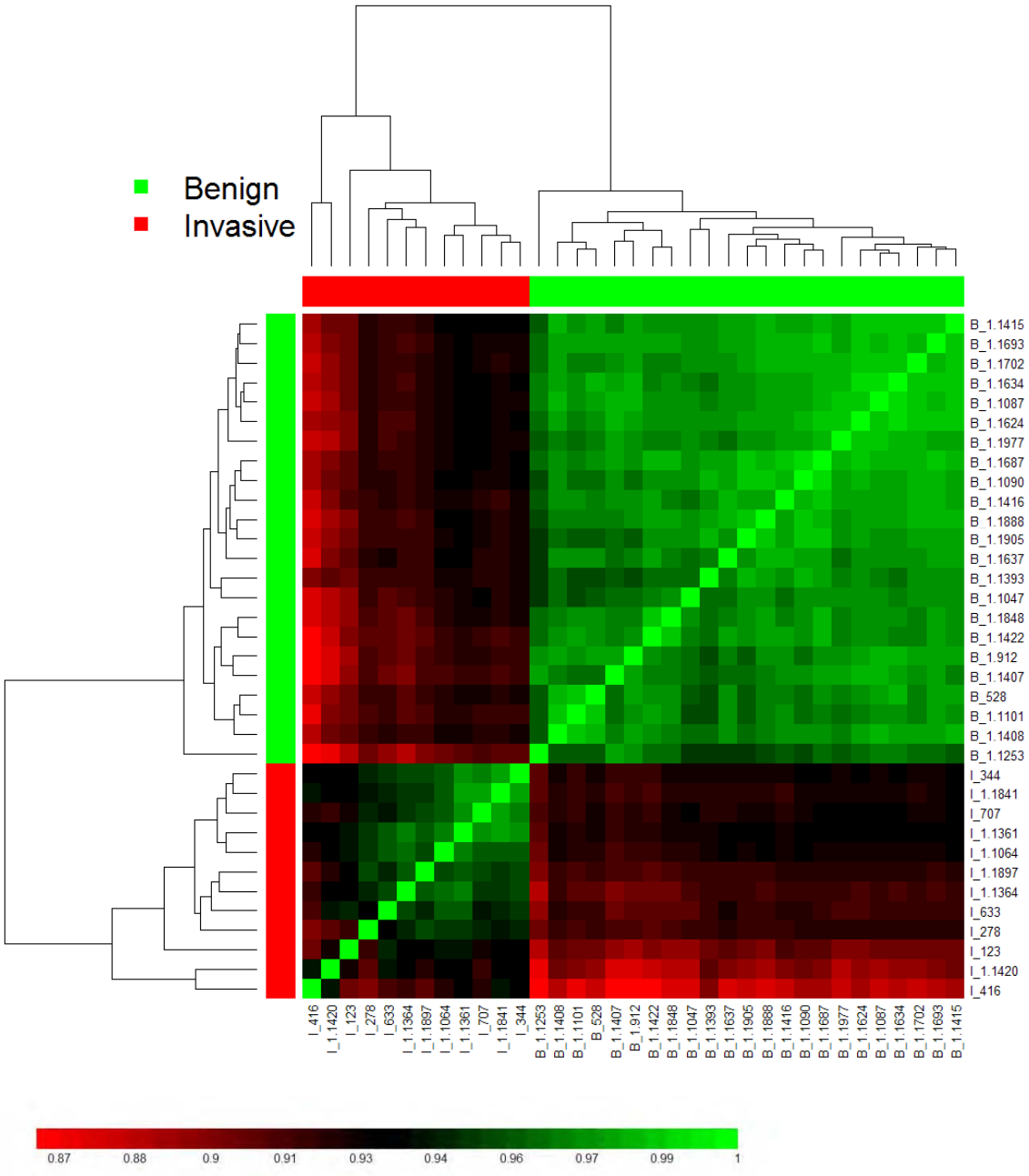
Boxplot:



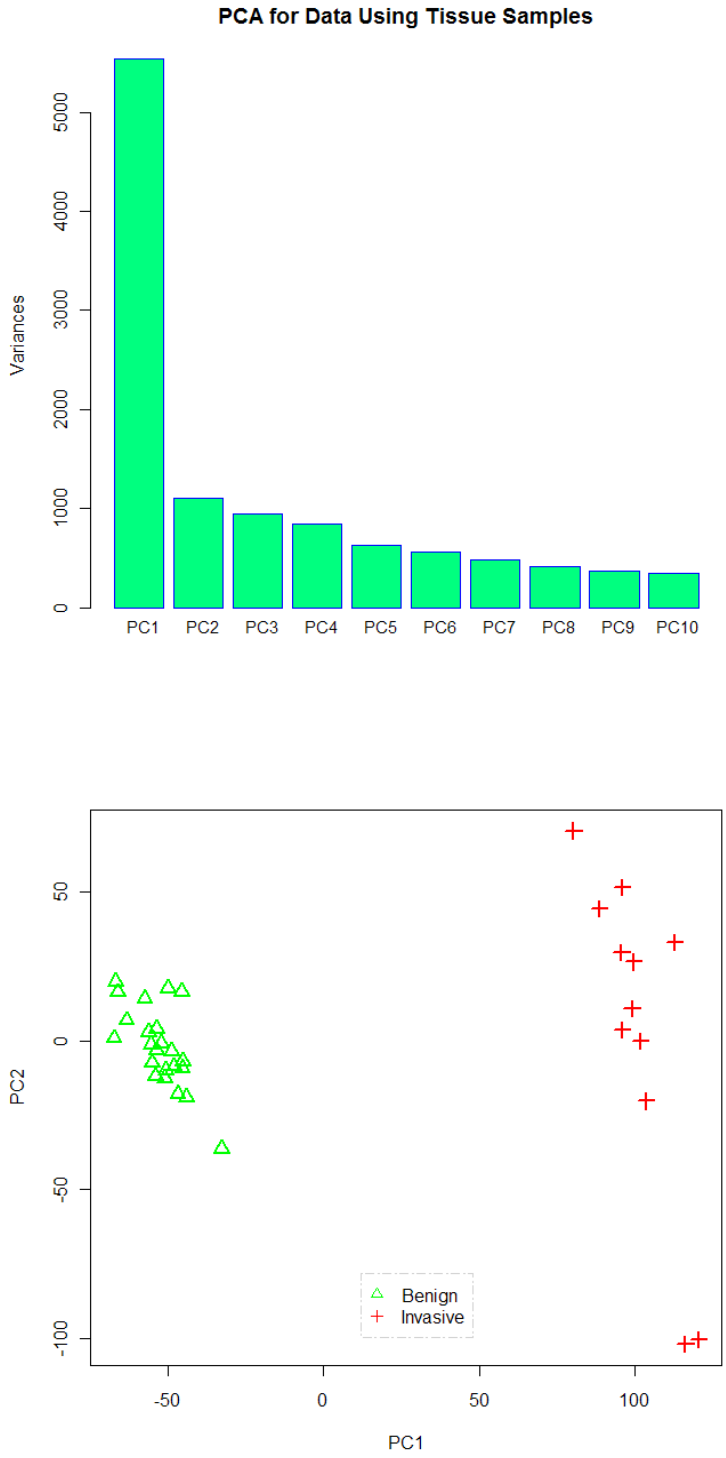
Clustering analysis of gene expression profiles:
The clustering results shows the benign and invasive groups are clustered together very well, with each group forming one cluster (see the graph below).



Visualization of correlations in samples:
Correlation analysis method is as described for the blood data. The heatmap and dendrogram visualization show that both invasive and benign group are cluster together as shown in the above hierarchical clustering analysis. Within the groups, samples have higher correlation comparing with between groups, which implies the gene expression profiles are similar. We can also see that the gene expression profiles are much better reflect the disease status of the samples than the ones using blood samples.



PCA analysis on microarray data using tissue sample:
We also carried out the PCA analysis for the microarray data using tissue samples. The result (see the figures below) shows that two groups are separately very well. The PC1 captures the majority varitaion between the two groups which is the biggest and also the most interestion variation. PC2 captures the most variation within the groups.



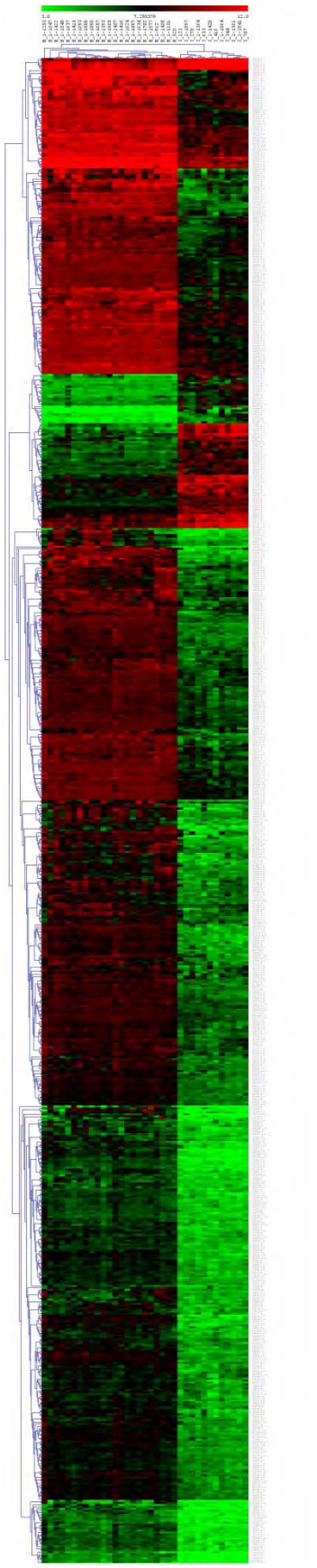
Diffrentailly gene expression patterns between benign and invasive:

Test: Wilcoxon Rank Sum Test
Significance Based on Est. FDR (Benjamini-Hochberg)
Selected FDR Limit: 1.0E-4
Computed FDR for Sig. Genes: 9.875525E-5

Group Information:
Benign (23 samples in analysis)
Invasive (12 samples in analysis)

Significant genes # of Significant Genes: 6091
% of Genes that are Signficant: 11%

Non-significant genes # of non-significant Genes: 48584
% of Genes that are not signficant: 89%
Among these 6091 transcripts, 1418 have FC greater than 3 and 763 with FC greater than 4. We can see most of the transcripts are changing in a small scale.
The follow heatmap visualizes the genes with FC > 4. For detail information about these genes, see the gene table.



Probesets	GENE_ TITLE	GENE_SYMBOL	ENTREZ_ID	p-value	adj.p (Benjamini/Hochberg)	AVE_B	AVE_I	FC
37892_at	collagen, type XI, alpha 1	COL11A1	1301	1.62E-06	4.16E-05	5.50	11.44	61.3
202037_s_at	secreted frizzled-related protein 1	SFRP1	6422	1.62E-06	4.16E-05	12.22	6.75	-44.6
217428_s_at	collagen, type X, alpha 1(Schmid metaphyseal chondrodysplasia) CO	L10A1	1300	1.62E-06	4.16E-05	4.67	10.14	44.2
1555229_a_at	complement component 1, s subcomponent C1S		716	1.62E-06	4.16E-05	9.30	3.86	-43.5
209560_s_at	delta-like 1 homolog (Drosophila) DLK1		8788	2.71E-06	4.39E-05	8.90	3.47	-43.0
204320_at	collagen, type XI, alpha 1	COL11A1	1301	1.62E-06	4.16E-05	5.16	10.47	39.5
217294_s_at eno	lase 1, (alpha)	ENO1	2023	1.62E-06	4.16E-05	10.56	5.36	-36.8
202035_s_at	secreted frizzled-related protein 1	SFRP1	6422	1.62E-06	4.16E-05	9.97	4.87	-34.1
204426_at	transmembrane emp24 domain trafficking protein 2	TMED2	10959	1.62E-06	4.16E-05	9.32	4.25	-33.8
213909_at	leucine rich repeat containing 15	LRRC15	131578	1.62 E-06	4.16E-05	6.15	11.15	31.9
202965_s_at cal	pain 6	CAPN6	827	1.62E-06	4.16E-05	9.02	4.21	-28.2
223623_at	chromosome 2 open reading frame 40	C2orf40	84417	1.62E-06	4.16E-05	10.36	5.55	-27.9
204427_s_at	transmembrane emp24 domain trafficking protein 2	TMED2 10	959	1.62E-06	4.16E-05	10.39	5.61	-27.6
202036_s_at	secreted frizzled-related protein 1	SFRP1	6422	1.62E-06	4.16E-05	10.32	5.54	-27.5
200869_at	ribosomal protein L18a | similar to ribosomal protein L18a; 60S ribosomal protein L18a	LOC390354 | RPL18A	390354 | 6142 1	.62E-06	4.16E-05	12.74	7.98	-27.0
205941_s_at	collagen, type X, alpha 1(Schmid metaphyseal chondrodysplasia)	COL10A1 13	00	1.62E-06	4.16E-05	5.49	10.23	26.6
214279_s_at	NDRG family member 2	NDRG2	57447 1	62E-06	4.16E-05	9.17	4.44	-26.5
206742_at	c-fos induced growth factor (vascular endothelial growth factor D)	FIGF	2277	1.62E-06	4.16E-05	9.30	4.66	-25.0
209613_s_at	alcohol dehydrogenase 1C (class I), gamma polypeptide | alcohol dehydrogenase IB (class I), beta polypeptide	ADH1B | ADH1C	125 | 126	1.62E-06	4.16E-05	9.52	4.89	-24.8
205044_at	gamma-aminobutyric acid (GABA) A receptor, pi	GABRP 2568		2.29E-06	4.19E-05 10.	52	5.91	-24.4
240724_at Tran	scribed locus	NA	NA	6.25E-06	6.66E-05	7.26	2.78	-22.4
204851_s_at	doublecortex; lissencephaly, X-linked (doublecortin)	DCX	1641	1.92E-06	4.16E-05	7.95	3.53	-21.3
204712_at	WNT inhibitory factor 1	WIF1	11197 4	49E-06	5.50E-05	9.15	4.74	-21.3
229479_at Tran	scribed locus	NA	NA	1.93E-06	4.16E-05	4.28	8.68	21.1
229839_at	Scavenger receptor class A, member 5 (putative)	SCARA5	286133	1.62E-06	4.16E-05	8.97	4.61	-20.6
1555814_a_at	ras homolog gene family, member A	RHOA 3	87	1.62E-06	4.16E-05	11.14	6.78	-20.5
209351_at	keratin 14 (epidermolysis bullosa simplex, Dowling-Meara, Koebner)	KRT14 38	61	1.93E-06	4.16E-05	10.28	5.97	-19.8
201367_s_at	zinc finger protein 36, C3H type-like 2	ZFP36L2	678	1.62E-06	4.16E-05	7.54	3.24	-19.7
1555730_a_at co	filin 1 (non-muscle)	CFL1	1072 1	.62E-06	4.16E-05	9.97	5.67	-19.7
222717_at	serum deprivation response (phosphatidylserine binding protein) SD	PR	8436	1.62E-06	4.16E-05	7.61	3.42	-18.2
221796_at	neurotrophic tyrosine kinase, receptor, type 2	NTRK2	4915	1.62E-06	4.16E-05	10.02	5.83	-18.2
223341_s_at	short coiled-coil protein	SCOC	60592 1	62E-06	4.16E-05	8.66	4.49	-18.1
205713_s_at	cartilage oligomeric matrix protein CO	MP	1311	1.93E-06	4.16E-05	5.84	10.00	17.9
210511_s_at i	nhibin, beta A	INHBA	3624	1.62E-06	4.16E-05	5.93	10.08	17.7
221618_s_at	TAF9B RNA polymerase II, TATA box binding protein	LOC728198 | TAF9B	51616 | 728198 1	.62E-06	4.16E-05	7.82	3.69	-17.5

	(TBP)-associated factor, 31kDa | similar to transcription associated factor 9B							
204455_at d	ystonin	DST	667	1.62E-06	4.16E-05	10.73	6.62	-17.3
209612_s_at	alcohol dehydrogenase 1C (class I), gamma polypeptide | alcohol dehydrogenase IB (class I), beta polypeptide	ADH1B | ADH1C	125 | 126	3.80E-06	5.03E-05	10.34	6.24	-17.1
1554411_at	catenin (cadherin-associated protein), beta 1, 88kDa	CTNNB1 14	99	1.62E-06	4.16E-05	9.80	5.76	-16.5
206825_at o	xytocin receptor	OXTR	5021	1.62E-06	4.16E-05	9.90	5.87	-16.3
209774_x_at	chemokine (C-X-C motif) ligand 2	CXCL2	2920	4.49E-06	5.50E-05	10.01	6.02	-15.8
207542_s_at	aquaporin 1 (Colton blood group) A	QP1	358	1.62E-06	4.16E-05	9.48	5.55	-15.3
201295_s_at NA		NA	NA	1.62E-06	4.16E-05	8.27	4.38	-14.8
232541_at	CDNA FLJ20099 fis, clone COL04544 NA		NA	1.93E-06	4.16E-05	8.29	4.40	-14.8
1567458_s_at	ras-related C3 botulinum toxin substrate 1 (rho family, small GTP binding protein Rac1)	RAC1 5	879	1.62E-06	4.16E-05	9.94	6.06	-14.8
1555725_a_at	regulator of G-protein signaling 5	RGS5	8490	1.62E-06	4.16E-05	9.32	5.44	-14.7
1558048_x_at NA		NA	NA	2.29E-06	4.19E-05	6.38	2.51	-14.6
202409_at	insulin- insulin-like growth factor 2 | insulin-like growth factor 2 (somatomedin A)	IGF2 | INS-IGF2	3481 | 723961	1.62E-06 4.	16E-05	11.12	7.26	-14.6
216918_s_at d	ystonin	DST	667	1.62E-06	4.16E-05	9.52	5.71	-14.0
215236_s_at	phosphatidylinositol binding clathrin assembly protein	PICALM	8301	1.62E-06	4.16E-05	8.01	4.21	-13.9
228851_s_at end	osulfine alpha	ENSA	2029	1.62E-06	4.16E-05	8.56	4.77	-13.9
215253_s_at	regulator of calcineurin 1	RCAN1	1827 1	.62E-06	4.16E-05	8.20	4.42	-13.7
209283_at	crystallin, alpha B	CRYAB	1410 1	62E-06	4.16E-05	10.76	6.99	-13.6
204850_s_at	doublecortex; lissencephaly, X-linked (doublecortin)	DCX	1641	3.79E-06	5.03E-05	7.45	3.68	-13.6
226237_at	MRNA full length insert cDNA clone EUROIMAGE 1913076	NA N	A	1.62E-06	4.16E-05	7.17	10.94	13.6
220037_s_at	lymphatic vessel endothelial hyaluronan receptor 1	LYVE1	10894	1.93E-06	4.16E-05	7.47	3.71	-13.5
227140_at	CDNA FLJ11041 fis, clone PLACE1004405 N	A	NA	1.62E-06	4.16E-05	7.01	10.76	13.4
211672_s_at	actin related protein 2/3 complex, subunit 4, 20kDa	ARPC4	10093	1.62E-06	4.16E-05	8.82	5.08	-13.4
202746_at	integral membrane protein 2A IT	M2A	9452	1.93E-06	4.16E-05	11.31	7.57	-13.3
229271_x_at	collagen, type XI, alpha 1	COL11A1	1301	1.62E-06	4.16E-05	2.43	6.15	13.1
210317_s_at	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, epsilon polypeptide	YWHAE 7	531	1.62E-06	4.16E-05	8.15	4.44	-13.1
212224_at	aldehyde dehydrogenase 1 family, member A1	ALDH1A1	216	2.71E-06	4.39E-05	9.85	6.15	-13.0
219795_at	solute carrier family 6 (amino acid transporter), member 14	SLC6A14	11254	5.30E-06	6.05E-05	7.99	4.30	-13.0
202274_at	actin, gamma 2, smooth muscle, enteric	ACTG2	72	1.62E-06	4.16E-05	10.18	6.49	-12.9
206163_at	mab-21-like 1 (C. elegans)	MAB21L1	4081 1	.62E-06	4.16E-05	9.00	5.30	-12.9
1555564_a_at	complement factor I	CFI	3426	1.93E-06	4.16E-05	8.21	4.54	-12.8
228399_at	odd-skipped related 1 (Drosophila)	OSR1 1	30497	1.62E-06	4.16E-05	7.48	3.82	-12.7
231258_at NA		NA	NA	1.62E-06	4.16E-05	7.55	3.89	-12.7
238018_at	hypothetical protein LOC285016 h	CG_1990170	285016	1.93E-06	4.16E-05	7.08	3.42	-12.7
231669_at	Selenoprotein P, plasma, 1	SEPP1	6414	1.62E-06	4.16E-05	8.23	4.60	-12.4
1554679_a_at	lysosomal associated protein transmembrane 4 beta	LAPTM4B 55	353	5.30E-06	6.05E-05	9.07	5.46	-12.2

218748_s_at	exocyst complex component 5 E	XOC5	10640	1.62E-06	4.16E-05	6.75	3.15	-12.2
203878_s_at	matrix metallopeptidase 11 (stromelysin 3)	MMP11	4320	1.62E-06	4.16E-05	7.09	10.69	12.1
236044_at	phosphatidic acid phosphatase type 2 domain containing 1A	PPAPDC1A 1	96051	2.71E-06	4.39E-05	4.55	8.14	12.1
205030_at	fatty acid binding protein 7, brain FA	BP7	2173	3.80E-06	5.03E-05	7.92	4.33	-12.0
201123_s_at	eukaryotic translation initiation factor 5A	EIF5A 1	984	1.62E-06	4.16E-05	10.78	7.19	-12.0
213456_at	sclerostin domain containing 1 SO	STDC1	25928	1.62E-06	4.16E-05	8.11	4.53	-12.0
213470_s_at	heterogeneous nuclear ribonucleoprotein H1 (H)	HNRPH1	3187	1.62E-06	4.16E-05	8.60	5.02	-12.0
1555106_a_at	CTD (carboxy-terminal domain, RNA polymerase II, polypeptide A) small phosphatase like 2	CTDSPL2	51496	1.62E-06	4.16E-05	6.87	3.29	-11.9
203372_s_at	suppressor of cytokine signaling 2	SOCS2	8835	1.62E-06	4.16E-05	7.30	3.73	-11.9
209292_at	Inhibitor of DNA binding 4, dominant negative helix-loop-helix protein	ID4 34	00	3.80E-06	5.03E-05	8.76	5.20	-11.8
235476_at	tripartite motif-containing 59	TRIM59	286827	1.62E-06	4.16E-05	4.35	7.91	11.8
220624_s_at	E74-like factor 5 (ets domain transcription factor)	ELF5	2001	5.30E-06	6.05E-05	8.00	4.44	-11.8
209686_at	S100 calcium binding protein B	S100B	6285	1.62E-06	4.16E-05	8.03	4.47	-11.8
208399_s_at end	othelin 3	EDN3	1908	1.62E-06	4.16E-05	7.21	3.66	-11.7
202817_s_at	synovial sarcoma translocation, chromosome 18	SS18 67	60	1.62E-06	4.16E-05	9.02	5.47	-11.7
211559_s_at cy	clin G2	CCNG2	901	1.62E-06	4.16E-05	9.81	6.26	-11.7
210198_s_at	proteolipid protein 1 (Pelizaeus-Merzbacher disease, spastic paraplegia 2, uncomplicated) PL	P1	5354	1.62E-06	4.16E-05	7.78	4.23	-11.7
201971_s_at	ATPase, H+ transporting, lysosomal 70kDa, V1 subunit A	ATP6V1A	523	1.62E-06	4.16E-05	7.89	4.35	-11.6
1554433_a_at	zinc finger protein 146	ZNF146	7705 1	.62E-06	4.16E-05	8.12	4.59	-11.6
211726_s_at	flavin containing monooxygenase 2 (non-functional)	FMO2 2	327	1.62E-06	4.16E-05	9.30	5.77	-11.5
201551_s_at	lysosomal-associated membrane protein 1	LAMP1	3916	1.62E-06	4.16E-05	9.01	5.51	-11.3
210875_s_at	zinc finger E-box binding homeobox 1	ZEB1	6935	1.62E-06	4.16E-05	6.61	3.12	-11.2
201820_at	keratin 5 (epidermolysis bullosa simplex, Dowling-Meara/Kobner/Weber-Cockayne types)	KRT5 3	852	6.25E-06	6.66E-05	9.36	5.87	-11.2
1555154_a_at	quaking homolog, KH domain RNA binding (mouse)	QKI 9	444	1.62E-06	4.16E-05	8.12	4.63	-11.2
213406_at	WD repeat and SOCS box-containing 1	WSB1	26118	1.62E-06	4.16E-05	8.07	4.59	-11.1
210839_s_at	ectonucleotide pyrophosphatase/phosphodiesterase 2 (autotaxin)	ENPP2	5168	1.62E-06	4.16E-05	9.19	5.72	-11.1
204475_at	matrix metallopeptidase 1 (interstitial collagenase)	MMP1	4312	2.29E-06 4	19E-05	3.30	6.77	11.1
227875_at	kelch-like 13 (Drosophila)	KLHL13	90293 1	62E-06	4.16E-05	9.36	5.90	-11.0
1555233_at	ras homolog gene family, member J	RHOJ	57381	1.62E-06	4.16E-05	7.74	4.29	-10.9
214040_s_at	gelsolin (amyloidosis, Finnish type)	GSN	2934	1.62E-06	4.16E-05	8.66	5.21	-10.9
235849_at	scavenger receptor class A, member 5 (putative)	SCARA5	286133	1.62E-06	4.16E-05	8.95	5.51	-10.9
208719_s_at	DEAD (Asp-Glu-Ala-Asp) box polypeptide 17	DDX17	10521	1.62E-06	4.16E-05	7.66	4.22	-10.8
1558093_s_at	matrin 3 | similar to Matrin-3 (Nuclear scaffold protein P130/MAT3)	LOC727839 | MATR3	727839 | 9782	1.62E-06 4	16E-05	8.51	5.09	-10.7

215707_s_at	prion protein (p27-30) (Creutzfeldt-Jakob disease, Gerstmann-Strausler-Scheinker syndrome, fatal familial insomnia)	PRNP	5621	1.62E-06	4.16E-05	9.15	5.73	-10.7
219436_s_at en	domucin	EMCN	51705	1.62E-06	4.16E-05	8.43	5.02	-10.7
222846_at	RAB8B, member RAS oncogene family	RAB8B	51762	1.62E-06	4.16E-05	7.62	4.21	-10.7
219059_s_at	lymphatic vessel endothelial hyaluronan receptor 1	LYVE1	10894	1.62E-06	4.16E-05	8.26	4.87	-10.5
214505_s_at	four and a half LIM domains 1 FH	L1	2273	1.62E-06	4.16E-05	8.96	5.58	-10.4
205782_at	fibroblast growth factor 7 (keratinocyte growth factor)	FGF7	2252	1.62E-06	4.16E-05	8.40	5.03	-10.4
205029_s_at	fatty acid binding protein 7, brain FA	BP7	2173	3.21E-06	4.66E-05	6.58	3.21	-10.3
205392_s_at	chemokine (C-C motif) ligand 14 | chemokine (C-C motif) ligand 15	CCL14 | CCL15	6358 | 6359	2.71E-06 4	39E-05	9.23	5.89	-10.2
239672_at Tran	scribed locus	NA	NA	1.62E-06	4.16E-05	7.22	3.88	-10.1
215780_s_at	SET translocation (myeloid leukemia-associated) | SET translocation (myeloid leukemia-associated) pseudogene | similar to Protein SET (Phosphatase 2A inhibitor I2PP2A) (I-2PP2A) (Template-activating factor I) (TAF-I) (HLA-DR-associated protein II) (PHAPII) (Inhibitor of granzyme A-activated DNase) (IGAAD)	LOC389168 | SET | hCG_1644608	389168 | 6418 | 642869	1.62E-06 4	16E-05	10.96	7.63	-10.1
227062_at	trophoblast-derived noncoding RNA	TncRNA	283131	1.62E-06	4.16E-05	7.56	10.88	10.0
209821_at i	nterleukin 33	IL33	90865	1.62E-06	4.16E-05	7.82	4.51	-9.9
211737_x_at	pleiotrophin (heparin binding growth factor 8, neurite growth-promoting factor 1)	PTN	5764	5.30E-06	6.05E-05	11.59	8.31	-9.7
230472_at	iroquois homeobox 1	IRX1	79192	1.62E-06	4.16E-05	8.63	5.35	-9.7
215719_x_at	Fas (TNF receptor superfamily, member 6)	FAS	355	1.93E-06	4.16E-05	7.29	4.01	-9.7
211450_s_at	mutS homolog 6 (E. coli)	MSH6	2956	1.62E-06	4.16E-05	7.18	3.90	-9.7
214359_s_at	heat shock protein 90kDa alpha (cytosolic), class B member 1	HSP90AB1 3	326	1.62E-06	4.16E-05	11.01	7.74	-9.6
200641_s_at	similar to tyrosine 3/tryptophan 5 - monooxygenase activation protein, zeta polypeptide | tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide	LOC650083 | YWHAZ	650083 | 7534	1.62E-06 4	.16E-05	11.00	7.74	-9.6
221423_s_at	Yip1 domain family, member 5	YIPF5	81555	1.62E-06	4.16E-05	8.91	5.65	-9.6
213901_x_at RNA	binding motif protein 9	RBM9	23543	1.62E-06 4	.16E-05	8.60	5.34	-9.6
220425_x_at	ropporin, rhophilin associated protein 1B	ROPN1B 1	52015	1.62E-06	4.16E-05	8.02	4.77	-9.5
1553685_s_at Sp	1 transcription factor	SP1	6667	1.62E-06	4.16E-05	6.82	3.58	-9.4
220735_s_at	SUMO1/sentrin specific peptidase 7	SENP7	57337	1.62E-06	4.16E-05	6.41	3.17	-9.4
206201_s_at m	esenchyme homeobox 2	MEOX2	4223	2.29E-06	4.19E-05	7.58	4.35	-9.4
232584_at	CDNA FLJ12328 fis, clone MAMMA1002145	NA NA		1.62E-06	4.16E-05	8.77	5.55	-9.3
1554726_at	zinc finger protein 655	ZNF655	79027 1	.62E-06	4.16E-05	7.86	4.64	-9.3
210170_at	PDZ and LIM domain 3	PDLIM3	27295	8.65E-06	8.19E-05	7.50	4.28	-9.3
235236_at	CDNA FLJ31436 fis, clone NT2NE2000636 NA		NA	1.62E-06	4.16E-05	7.42	4.21	-9.2
203400_s_at t	ransferrin	TF	7018	1.62E-06	4.16E-05	8.73	5.53	-9.2

203434_s_at	membrane metallo-endopeptidase MME		4311	3.21E-06	4.66E-05	8.65	5.45	-9.1
209047_at	aquaporin 1 (Colton blood group) A	QP1	358	1.62E-06	4.16E-05	10.44	7.25	-9.1
201890_at	ribonucleotide reductase M2 polypeptide RRM2		6241	5.30E-06	6.05E-05	6.19	9.38	9.1
209170_s_at g	lycoprotein M6B	GPM6B	2824	1.93E-06	4.16E-05	9.97	6.78	-9.1
202747_s_at	integral membrane protein 2A	ITM2A 9	452	2.29E-06	4.19E-05	11.37	8.18	-9.1
222458_s_at	chromosome 1 open reading frame 108	C1orf108	79647	1.62E-06	4.16E-05	7.71	4.54	-9.0
218002_s_at	chemokine (C-X-C motif) ligand 14	CXCL14	9547	1.02E-05	9.16E-05	11.88	8.71	-9.0
1553105_s_at d	esmoglein 2	DSG2	1829	1.62E-06	4.16E-05	8.25	5.09	-9.0
209466_x_at	pleiotrophin (heparin binding growth factor 8, neurite growth-promoting factor 1)	PTN	5764	4.49E-06	5.50E-05	11.00	7.84	-8.9
228653_at	sterile alpha motif domain containing 5	SAMD5	389432	6.25E-06	6.66E-05	8.86	5.70	-8.9
1553749_at	family with sequence similarity 76, member B	FAM76B	143684	1.62E-06	4.16E-05	9.00	5.84	-8.9
217301_x_at	retinoblastoma binding protein 4	RBBP4 5928		1.62E-06	4.16E-05	10. 20	7.06	-8.8
217504_at	ATP-binding cassette, sub-family A (ABC1), member 6	ABCA6	23460	1.62E-06	4.16E-05	7.32	4.20	-8.7
201337_s_at	vesicle-associated membrane protein 3 (cellubrevin)	VAMP3	9341	1.62E-06	4.16E-05	9.16	6.04	-8.7
1558101_at NA		NA	NA	1.62E-06	4.16E-05	7.46	4.34	-8.7
206544_x_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	SMARCA2	6595	1.62E-06	4.16E-05	9.26	6.15	-8.6
209541_at	insulin-like growth factor 1 (somatomedin C)	IGF1	3479	1.62E-06	4.16E-05	11.04	7.93	-8.6
208047_s_at	NGFI-A binding protein 1 (EGR1 binding protein 1)	NAB1	4664	1.62E-06	4.16E-05	6.14	3.04	-8.6
228268_at	flavin containing monooxygenase 2 (non-functional) FMO2		2327	7.36E-06	7.39E-05	10.06	6.95	-8.6
204469_at	protein tyrosine phosphatase, receptor-type, Z polypeptide 1 PTPRZ1		5803	1.62E-06	4.16E-05	6.50	3.41	-8.5
204748_at	prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase) PT	GS2	5743	7.36E-06	7.39E-05	8.01	4.92	-8.5
219929_s_at	zinc finger, FYVE domain containing 21	ZFYVE21	79038	1.62E-06	4.16E-05	7.75	4.66	-8.5
210298_x_at	four and a half LIM domains 1 FHL1		2273	1.62E-06	4.16E-05	8.02	4.94	-8.5
241705_at	ATP-binding cassette, sub-family A (ABC1), member 5	ABCA5	23461	2.29E-06	4.19E-05	8.10	5.02	-8.5
233303_at	Homo sapiens, clone IMAGE:4295366, mRNA	NA	NA	1.62E-06	4.16E-05	8.52	5.44	-8.4
1552509_a_at	CD300 molecule-like family member g	CD300LG	146894	2.28E-06	4.19E-05	8.41	5.33	-8.4
211467_s_at	nuclear factor I/B	NFIB	4781	1.62E-06	4.16E-05	8.60	5.53	-8.4
216252_x_at	Fas (TNF receptor superfamily, member 6)	FAS	355	1.93E-06	4.16E-05	7.29	4.21	-8.4
209763_at cho	rdin-like 1	CHRDL1	91851	4.49E-06	5.50E-05	9.96	6.89	-8.4
1555745_a_at	lysozyme (renal amyloidosis) LYZ		4069	1.93E-06	4.16E-05	7.51	4.45	-8.3
1564494_s_at	procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), beta polypeptide P4	HB	5034	1.93E-06	4.16E-05	8.52	5.47	-8.3
1556579_s_at	immunoglobulin superfamily, member 10	IGSF10	285313	1.62E-06	4.16E-05	7.73	4.68	-8.3
204273_at	endothelin receptor type B	EDNRB	1910	1.62E-06	4.16E-05	8.34	5.29	-8.3
205018_s_at	muscleblind-like 2 (Drosophila) MBNL2		10150	1.62E-06	4.16E-05	8.43	5.38	-8.3
230867_at	collagen type VI alpha 6	COL6A6	131873	3.80E-06	5.03E-05	7.68	4.63	-8.3

221795_at	neurotrophic tyrosine kinase, receptor, type 2	NTRK2	4915	1.62E-06	4.16E-05	7.34	4.29	-8.2
224221_s_at	vav 3 guanine nucleotide exchange factor	VAV3	10451	1.62E-06	4.16E-05	6.68	3.64	-8.2
214449_s_at	ras homolog gene family, member Q	RHOQ	23433	1.62E-06	4.16E-05	8.65	5.62	-8.2
1555240_s_at	guanine nucleotide binding protein (G protein), gamma 12	GNG12 55	970	1.62E-06	4.16E-05	7.81	4.78	-8.2
210299_s_at	four and a half LIM domains 1	FHL1 2	273	1.62E-06	4.16E-05	7.98	4.95	-8.2
236028_at	Integrin-binding sialoprotein (bone sialoprotein, bone sialoprotein II)	IBSP 3	381	1.93E-06	4.16E-05	2.64	5.66	8.1
1553694_a_at	phosphoinositide-3-kinase, class 2, alpha polypeptide	PIK3C2A	5286	1.62E-06	4.16E-05	7.61	4.59	-8.1
219872_at	chromosome 4 open reading frame 18	C4orf18 51	313	1.62E-06	4.16E-05	7.41	4.39	-8.1
211317_s_at	CASP8 and FADD-like apoptosis regulator	CFLAR	8837	1.62E-06	4.16E-05	8.55	5.54	-8.1
231535_x_at	ropporin, rhophilin associated protein 1	ROPN1	54763	1.62E-06	4.16E-05	7.95	4.94	-8.0
222242_s_at	kallikrein-related peptidase 5	KLK5	25818	1.93E-06	4.16E-05	8.38	5.37	-8.0
202410_x_at	insulin- insulin-like growth factor 2 | insulin-like growth factor 2 (somatomedin A)	IGF2 | INS-IGF2	3481 | 723961	3.21E-06 4	.66E-05	7.94	4.94	-8.0
1554464_a_at carti	lage associated protein	CRTAP 1	0491	1.62E-06	4.16E-05	9.03	6.03	-8.0
200008_s_at	GDP dissociation inhibitor 2	GDI2	2665	1.62E-06	4.16E-05	10.60	7.60	-8.0
203088_at fi	bulin 5	FBLN5	10516	1.62E-06	4.16E-05	9.77	6.77	-8.0
214544_s_at	synaptosomal-associated protein, 23kDa	SNAP23	8773	1.62E-06	4.16E-05	8.45	5.46	-8.0
209458_x_at	hemoglobin, alpha 1 | hemoglobin, alpha 2	HBA1 | HBA2	3039 | 3040 7	.36E-06	7.39E-05	9.17	6.18	-8.0
203032_s_at fu	marate hydratase	FH	2271	1.62E-06	4.16E-05	6.36	3.37	-7.9
220606_s_at	chromosome 17 open reading frame 48	C17orf48	56985	1.62E-06	4.16E-05	7.50	4.52	-7.9
205528_s_at	runt-related transcription factor 1; translocated to, 1 (cyclin D-related)	RUNX1T1	862	1.62E-06	4.16E-05	9.01	6.03	-7.9
205321_at	eukaryotic translation initiation factor 2, subunit 3 gamma, 52kDa	EIF2S3 1	968	1.62E-06	4.16E-05	10.58	7.60	-7.9
235388_at	chromodomain helicase DNA binding protein 9	CHD9	80205	1.62E-06	4.16E-05	7.79	4.81	-7.9
201718_s_at	erythrocyte membrane protein band 4.1-like 2	EPB41L2	2037	1.62E-06	4.16E-05	8.37	5.40	-7.9
238332_at	ankyrin repeat domain 29	ANKRD29	147463	1.62E-06	4.16E-05	7.44	4.47	-7.8
1558214_s_at	catenin (cadherin-associated protein), alpha 1, 102kDa	CTNNA1	1495	1.62E-06	4.16E-05	6.66	3.70	-7.8
1554868_s_at	PEST proteolytic signal containing nuclear protein	PCNP	57092	1.62E-06	4.16E-05	10.68	7.72	-7.8
235489_at	ras homolog gene family, member J	RHOJ	57381	1.62E-06	4.16E-05	8.89	5.93	-7.8
239849_at Tran	scribed locus	NA	NA	6.25E-06	6.66E-05	7.73	4.78	-7.7
212353_at su	lfatase 1	SULF1	23213	1.62E-06	4.16E-05	7.94	10.89	7.7
238906_s_at	ras homolog gene family, member J	RHOJ	57381	1.62E-06	4.16E-05	7.88	4.92	-7.7
221268_s_at	sphingosine-1-phosphate phosphatase 1	SGPP1	81537	1.92E-06	4.16E-05	7.91	4.96	-7.7
1555724_s_at t	ransgelin	TAGLN	6876	1.62E-06	4.16E-05	11.44	8.49	-7.7
206767_at	RNA binding motif, single stranded interacting protein	RBMS3	27303	1.62E-06	4.16E-05	8.21	5.26	-7.7
1559881_s_at	zinc finger protein 12	ZNF12	7559 1	.62E-06	4.16E-05	8.10	5.16	-7.7
202966_at cal	pain 6	CAPN6	827	1.62E-06	4.16E-05	6.97	4.03	-7.7
238461_at	eukaryotic translation initiation factor 4E family member 3	EIF4E3 3	17649	1.62E-06	4.16E-05	7.10	4.16	-7.7
242518_at	CDNA FLJ43403 fis, clone OCBBF2016612 NA		NA	1.62E-06	4.16 E-05	6.92	3.99	-7.7
1555460_a_at	solute carrier family 39 (zinc transporter), member 6	SLC39A6	25800	4.49E-06	5.50E-05	9.59	6.66	-7.6

201120_s_at	progesterone receptor membrane component 1	PGRMC1	10857	1.62E-06	4.16E-05	10.20	7.27	-7.6
205960_at	pyruvate dehydrogenase kinase, isozyme 4	PDK4	5166	1.62E-06	4.16E-05	7.07	4.14	-7.6
210458_s_at	TRAF family member-associated NFKB activator	TANK	10010	1.62E-06	4.16E-05	6.82	3.90	-7.6
209540_at	insulin-like growth factor 1 (somatomedin C)	IGF1	3479	5.30E-06	6.05E-05	10.62	7.70	-7.5
228107_at NA		NA	NA	1.62E-06	4.16E-05	6.78	3.86	-7.5
209189_at	v-fos FBJ murine osteosarcoma viral oncogene homolog	FOS 2	353	6.25E-06	6.66E-05	11.56	8.64	-7.5
1570507_at	Splicing factor, arginine/serine-rich 2, interacting protein	SFRS2IP 9	169	1.62E-06	4.16E-05	7.19	4.28	-7.5
209318_x_at	pleiomorphic adenoma gene-like 1	PLAGL1	5325	1.62E-06	4.16E-05	9.62	6.71	-7.5
219932_at	solute carrier family 27 (fatty acid transporter), member 6	SLC27A6	28965	1.62E-06	4.16E-05	7.20	4.30	-7.5
225575_at	leukemia inhibitory factor receptor alpha	LIFR	3977	1.62E-06	4.16E-05	10.14	7.24	-7.5
209629_s_at	nuclear transport factor 2-like export factor 2	NXT2	55916	1.62E-06	4.16E-05	7.14	4.25	-7.4
217811_at s	elenoprotein T	SELT	51714	1.93E-06	4.16E-05	10.14	7.24	-7.4
220477_s_at	chromosome 20 open reading frame 30	C20orf30	29058	1.62E-06	4.16E-05	9.60	6.72	-7.4
210145_at	phospholipase A2, group IVA (cytosolic, calcium-dependent)	PLA2G4A 5	321	5.30E-06	6.05E-05	8.12	5.23	-7.4
202504_at	tripartite motif-containing 29	TRIM29	23650	4.87E-06	5.94E-05	9.67	6.79	-7.4
1554614_a_at	polypyrimidine tract binding protein 2	PTBP2	58155	1.62E-06	4.16E-05	6.97	4.09	-7.4
206113_s_at	RAB5A, member RAS oncogene family	RAB5A	5868	1.62E-06	4.16E-05	8.27	5.40	-7.3
239512_at	splicing factor, arginine/serine-rich 4	SFRS4	6429	1.62E-06	4.16E-05	7.81	4.94	-7.3
232165_at ep	iplakin 1	EPPK1	83481	1.93E-06	4.16E-05	7.29	10.16	7.3
203108_at	G protein-coupled receptor, family C, group 5, member A GPRC5	A	9052	2.29E-06	4.19E-05	7.05	9.92	7.3
211699_x_at	hemoglobin, alpha 1 | hemoglobin, alpha 2	HBA1 | HBA2	3039 | 3040 8	.65E-06	8.19E-05	9.10	6.24	-7.3
212942_s_at K	IAA1199	KIAA1199	57214	3.21E-06	4.66E-05	4.47	7.31	7.2
238617_at	CDNA FLJ38181 fis, clone FCBBF1000125 NA		NA	3.21E-06	4.66E-05	5.79	8.63	7.2
229802_at	CDNA FLJ14388 fis, clone HEMBA1002716 N	A	NA	7.36E-06	7.39E-05	6.63	9.47	7.2
1556834_at	CDNA clone IMAGE:5296106 NA		NA	1.62E-06	4.16E-05	6.07	3.23	-7.2
224191_x_at	ropporin, rhophilin associated protein 1	ROPN1 54	763	1.62E-06	4.16E-05	7.90	5.06	-7.2
227850_x_at	CDC42 effector protein (Rho GTPase binding) 5	CDC42EP5	148170 2	.71E-06	4.39E-05	7.96	5.13	-7.1
226933_s_at	Inhibitor of DNA binding 4, dominant negative helix-loop-helix protein	ID4 3	400	3.79E-06	5.03E-05	8.12	5.29	-7.1
211734_s_at	Fc fragment of IgE, high affinity I, receptor for; alpha polypeptide	FCER1A 2	205	6.25E-06	6.66E-05	8.08	5.25	-7.1
203387_s_at	TBC1 domain family, member 4	TBC1D4 9	882	1.62E-06	4.16E-05	9.36	6.54	-7.1
215913_s_at	GULP, engulfment adaptor PTB domain containing 1	GULP1	51454	2.29E-06	4.19E-05	7.74	4.91	-7.1
200672_x_at	spectrin, beta, non-erythrocytic 1	SPTBN1	6711	1.62E-06	4.16E-05	9.45	6.63	-7.1
221430_s_at	ring finger protein 146	RNF146	81847	1.62E-06	4.16E-05	8.20	5.38	-7.1
209167_at g	lycoprotein M6B	GPM6B	2824	2.71E-06	4.39E-05	10.34	7.52	-7.1
225681_at	collagen triple helix repeat containing 1	CTHRC1	115908	5.30E-06	6.05E-05	9.68	12.50	7.0
209168_at g	lycoprotein M6B	GPM6B	2824	1.91E-06	4.16E-05	8.61	5.80	-7.0
239568_at	pleckstrin homology domain containing, family H (with	PLEKHH2 1	30271	2.29E-06	4.19E-05	6.65	3.83	-7.0

	MyTH4 domain) member 2							
202137_s_at	zinc finger, MYND domain containing 11	ZMYND11	10771	1.62E-06	4.16E-05	8.39	5.58	-7.0
222573_s_at	salvador homolog 1 (Drosophila) SAV1		60485	1.62E-06	4.16E-05	9.46	6.65	-7.0
229201_at	Full-length cDNA clone CS0DF014YC15 of Fetal brain of Homo sapiens (human)	NA NA		1.62E-06	4.16E-05	6.59	3.79	-7.0
206336_at	chemokine (C-X-C motif) ligand 6 (granulocyte chemotactic protein 2)	CXCL6 6	372	8.65E-06	8.19E-05	6.33	3.52	-7.0
201539_s_at	four and a half LIM domains 1 FHL1		2273	1.62E-06	4.16E-05	7.56	4.76	-7.0
1553725_s_at	zinc finger protein 644	ZNF644	84146 1	.62E-06	4.16E-05	7.37	4.56	-7.0
216899_s_at	src kinase associated phosphoprotein 2	SKAP2	8935	1.62E-06	4.16E-05	7.94	5.14	-7.0
242541_at	ATP-binding cassette, sub-family A (ABC1), member 9	ABCA9	10350	1.62E-06	4.16E-05	7.07	4.28	-6.9
210729_at	neuropeptide Y receptor Y2	NPY2R	4887	7.97E-06	7.98E-05	5.43	2.64	-6.9
204456_s_at g	rowth arrest-specific 1	GAS1	2619	1.93E-06	4.16E-05	7.04	4.25	-6.9
214581_x_at	tumor necrosis factor receptor superfamily, member 21	TNFRSF21 27	242	5.30E-06	6.05E-05	7.98	5.19	-6.9
209465_x_at	pleiotrophin (heparin binding growth factor 8, neurite growth-promoting factor 1)	PTN 5	764	7.36E-06	7.39E-05	9.60	6.81	-6.9
217300_at NA		NA	NA	1.62E-06	4.16E-05	8.16	5.37	-6.9
221428_s_at	transducin (beta)-like 1X-linked receptor 1	TBL1XR1	79718	1.62E-06	4.16E-05	9.44	6.67	-6.9
203951_at	calponin 1, basic, smooth muscle CNN1		1264	1.62E-06	4.16E-05	8.90	6.12	-6.8
211376_s_at	non-SMC element 4 homolog A (S. cerevisiae)	NSMCE4A	54780	1.62E-06	4.16E-05	9.31	6.54	-6.8
212444_at	CDNA clone IMAGE:6025865	NA NA		3.21E-06	4.66E-05	7.20	9.97	6.8
1554678_s_at	heterogeneous nuclear ribonucleoprotein D-like	HNRPDL 9	987	1.62E-06 4	.16E-05	9.59	6.82	-6.8
224173_s_at	mitochondrial ribosomal protein L30	MRPL30	51263	1.62E-06	4.16E-05	6.35	3.59	-6.8
228729_at cy	clin B1	CCNB1	891	1.62E-06	4.16E-05	5.32	8.07	6.7
208983_s_at	platelet/endothelial cell adhesion molecule (CD31 antigen) PECAM1		5175	1.62E-06	4.16E-05	8.41	5.66	-6.7
229127_at	CDNA FLJ31517 fis, clone NT2RI2000007 NA		NA	1.62E-06	4.16E-05	9.05	6.31	-6.7
225687_at	family with sequence similarity 83, member D	FAM83D	81610	5.30E-06	6.05E-05	5.02	7.76	6.7
219564_at	potassium inwardly-rectifying channel, subfamily J, member 16	KCNJ16	3773	1.62E-06	4.16E-05	6.17	3.44	-6.7
227404_s_at	Early growth response 1	EGR1	1958	2.71E-06	4.39E-05	11.41	8.67	-6.7
204969_s_at radi	xin	RDX	5962	1.62E-06	4.16E-05	6.25	3.52	-6.6
200082_s_at	ribosomal protein S7 | similar to 40S ribosomal protein S7 (S8)	LOC644315 | RPS7	6201 | 644315	1.62E-06 4	.16E-05	12.54	9.81	-6.6
213258_at	tissue factor pathway inhibitor (lipoprotein-associated coagulation inhibitor)	TFPI 7	035	4.49E-06	5.50E-05	10.15	7.44	-6.6
200634_at p	rofilin 1	PFN1	5216	1.62E-06	4.16E-05	10.46	7.74	-6.6
217762_s_at	RAB31, member RAS oncogene family	RAB31 11	031	1.62E-06	4.16E-05	8.77	11.49	6.6
210904_s_at	interleukin 13 receptor, alpha 1	IL13RA1	3597	1.62E-06	4.16E-05	8.69	5.97	-6.6
206701_x_at	endothelin receptor type B	EDNRB	1910	1.62E-06	4.16E-05	8.26	5.55	-6.6
209842_at	SRY (sex determining region Y)-box 10	SOX10	6663	1.62E-06	4.16E-05	7.97	5.26	-6.5
201308_s_at s	eptin 11	11-Sep	55752	1.62E-06	4.16E-05	6.95	4.25	-6.5
219679_s_at	WW domain containing	WAC	51322	1.62E-06	4.16E-05	8.71	6.02	-6.5

	adaptor with coiled-coil							
213415_at	chloride intracellular channel 2	CLIC2	1193	1.62E-06	4.16E-05	8.11	5.42	-6.5
1558014_s_at	male sterility domain containing 2	MLSTD2	84188	1.62E-06	4.16E-05	6.09	3.40	-6.5
1555609_a_at	zinc finger, matrin type 3	ZMAT3 64	393	1.62E-06	4.16E-05	6.84	4.15	-6.5
233608_at	CDNA FLJ11929 fis, clone HEMBB1000434 NA		NA	1.62E-06	4.16E-05	6.97	4.28	-6.4
1558199_at fi	bronectin 1	FN1	2335	6.25E-06	6.66E-05	5.62	8.31	6.4
202342_s_at	tripartite motif-containing 2	TRIM2	23321	3.21E-06	4.66E-05	9.23	6.54	-6.4
228186_s_at	R-spondin 3 homolog (Xenopus laevis)	RSPO3	84870	1.93E-06	4.16E-05	8.60	5.92	-6.4
218706_s_at	GRAM domain containing 3	GRAMD3	65983	1.62E-06	4.16E-05	10.39	7.71	-6.4
201742_x_at	splicing factor, arginine/serine-rich 1 (splicing factor 2, alternate splicing factor)	SFRS1	6426	1.62E-06	4.16E-05	10.09	7.41	-6.4
228434_at b	utyrophilin-like 9	BTNL9	153579	1.62E-06	4.16E-05	9.26	6.58	-6.4
209392_at	ectonucleotide pyrophosphatase/phosphodie sterase 2 (autotaxin)	ENPP2	5168	1.62E-06	4.16E-05	10.69	8.01	-6.4
207426_s_at	tumor necrosis factor (ligand) superfamily, member 4 (tax-transcriptionally activated glycoprotein 1, 34kDa)	TNFSF4 7292		5.30E-06	6.05E-05	4.35	7.02	6.4
217494_s_at	phosphatase and tensin homolog (mutated in multiple advanced cancers 1), pseudogene 1	PTENP1	11191	1.62E-06	4.16E-05	6.05	3.38	-6.4
201508_at	insulin-like growth factor binding protein 4	IGFBP4	3487	1.62E-06	4.16E-05	10.74	8.08	-6.3
235759_at Tran	scribed locus	NA	NA	1.62E-06	4.16E-05	7.37	4.70	-6.3
209535_s_at NA		NA	NA	2.29E-06	4.19E-05	7.43	4.77	-6.3
212730_at desm	uslin	DMN	23336	2.71E-06	4.39E-05	11.14	8.48	-6.3
210425_x_at	golgi autoantigen, golgin subfamily a, 8B	GOLGA8B	440270	1.62E-06	4.16E-05	9.91	7.25	-6.3
207980_s_at	Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 2	CITED2 10	370	1.62E-06	4.16E-05	7.93	5.27	-6.3
214895_s_at	ADAM metallopeptidase domain 10	ADAM10	102	1.62E-06	4.16E-05	7.63	4.97	-6.3
229568_at	MOB1, Mps One Binder kinase activator-like 2B (yeast) MOBKL2	B	79817	1.62E-06	4.16E-05	7.92	5.26	-6.3
227719_at Tran	scribed locus	NA	NA	1.62E-06	4.16E-05	8.29	5.64	-6.3
238010_at	chromosome 1 open reading frame 174	C1orf174	339448	1.62E-06	4.16E-05	7.19	4.54	-6.3
201950_x_at	capping protein (actin filament) muscle Z-line, beta	CAPZB	832	1.62E-06	4.16E-05	9.95	7.29	-6.3
205893_at n	euroligin 1	NLGN1	22871	2.71E-06	4.39E-05	5.87	3.22	-6.3
228195_at	hypothetical protein MGC13057 MGC13	057	84281	1.62E-06	4.16E-05	7.29	4.64	-6.3
1558015_s_at	ARP2 actin-related protein 2 homolog (yeast)	ACTR2	10097	1.62E-06	4.16E-05	8.48	5.83	-6.3
227709_at	Pp13759 | Similar to Reticulocalbin-1 precursor	LOC728913 7	28913	1.62E-06	4.16E-05	8.16	5.51	-6.3
214078_at	Primary neuroblastoma cDNA, clone:Nbla04246, full insert sequence	NA NA		1.62E-06	4.16E-05	7.24	4.59	-6.3
201130_s_at	cadherin 1, type 1, E-cadherin (epithelial)	CDH1	999	3.21E-06	4.66E-05	7.95	5.30	-6.3
214063_s_at t	ransferrin	TF	7018	2.71E-06	4.39E-05	8.48	5.83	-6.3
219326_s_at	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 2	B3GNT2	10678	1.62E-06	4.16E-05	8.43	5.79	-6.3
242137_at	CDNA FLJ36544 fis, clone TRACH2006378	NA NA		1.62E-06	4.16E-05	6.63	3.98	-6.3

233496_s_at co	filin 2 (muscle)	CFL2	1073	1.93E-06	4.16E-05	6.70	4.06	-6.2
211194_s_at	tumor protein p63	TP63	8626	1.62E-06	4.16E-05	6.65	4.01	-6.2
209863_s_at	tumor protein p63	TP63	8626	1.93E-06	4.16E-05	8.58	5.94	-6.2
203811_s_at	DnaJ (Hsp40) homolog, subfamily B, member 4	DNAJB4	11080	2.28E-06	4.19E-05	7.91	5.28	-6.2
215169_at	solute carrier family 35, member E2	SLC35E2	9906	3.21E-06	4.66E-05	5.85	8.48	6.2
231762_at	fibroblast growth factor 10	FGF10	2255	1.92E-06	4.16E-05	6.73	4.10	-6.2
206091_at mat	rilin 3	MATN3	4148	3.80E-06	5.03E-05	3.59	6.22	6.2
201151_s_at	muscleblind-like (Drosophila) MBNL1		4154	1.62E-06	4.16E-05	8.62	5.99	-6.2
201116_s_at carbox	ypeptidase E	CPE	1363	1.62E-06	4.16E-05	11.34	8.71	-6.2
209993_at	ATP-binding cassette, sub-family B (MDR/TAP), member 1	ABCB1 5243		1.62E-06	4.16E-05	6.91	4.29	-6.2
201044_x_at	dual specificity phosphatase 1	DUSP1 1	843	3.21E-06	4.66E-05	7.98	5.36	-6.2
205991_s_at	paired related homeobox 1	PRRX1	5396	1.62E-06	4.16E-05	8.03	5.41	-6.1
234975_at	CDNA FLJ38048 fis, clone CTONG2014264 | CDNA FLJ39067 fis, clone NT2RP7014910	NA NA		1.62E-06	4.16E-05	6.77	4.16	-6.1
1559296_at	Hypothetical protein LOC730057 LOC7	30057	730057	1.62E-06	4.16E-05	6.64	4.02	-6.1
206667_s_at	secretory carrier membrane protein 1	SCAMP1 9	522	1.62E-06	4.16E-05	7.86	5.24	-6.1
204326_x_at met	allothionein 1X	MT1X	4501	4.49E-06	5.50E-05	11.17	8.55	-6.1
201058_s_at	myosin, light chain 9, regulatory MYL9		10398	1.93E-06	4.16E-05	9.62	7.00	-6.1
227183_at	CDNA FLJ36638 fis, clone TRACH2018950 NA		NA	1.62E-06	4.16E-05	8.66	6.05	-6.1
208960_s_at	Kruppel-like factor 6	KLF6	1316	1.62E-06	4.16E-05	9.09	6.48	-6.1
238918_at	CDNA FLJ42015 fis, clone SPLEN2032813 NA		NA	1.62E-06	4.16E-05	7.97	5.36	-6.1
211537_x_at	mitogen-activated protein kinase kinase kinase 7	MAP3K7	6885	1.62E-06	4.16E-05	7.75	5.14	-6.1
202226_s_at	v-crk sarcoma virus CT10 oncogene homolog (avian)	CRK 1	398	1.62E-06	4.16E-05	9.52	6.91	-6.1
241981_at	family with sequence similarity 20, member A	FAM20A	54757	2.29E-06	4.19E-05	7.15	4.55	-6.1
213087_s_at	CDNA clone IMAGE:4838699 NA		NA	1.62E-06	4.16E-05	7.02	4.42	-6.1
204011_at	sprouty homolog 2 (Drosophila) SPRY2		10253	1.62E-06	4.16E-05	9.23	6.63	-6.1
212398_at radi	xin	RDX	5962	1.62E-06	4.16E-05	7.43	4.83	-6.1
202431_s_at	v-myc myelocytomatosis viral oncogene homolog (avian)	MYC 4	609	1.93E-06	4.16E-05	10.97	8.38	-6.0
214701_s_at fi	bronectin 1	FN1	2335	1.62E-06	4.16E-05	5.78	8.36	6.0
1556285_s_at	pyrophosphatase (inorganic) 2	PPA2 27	068	1.62E-06	4.16E-05	9.89	7.31	-6.0
1555167_s_at	pre-B-cell colony enhancing factor 1	PBEF1	10135	2.29E-06	4.19E-05	7.40	4.82	-6.0
1553106_at	chromosome 5 open reading frame 24	C5orf24	134553	1.62E-06	4.16E-05	7.01	4.43	-6.0
205848_at g	rowth arrest-specific 2	GAS2	2620	1.93E-06	4.16E-05	6.84	4.27	-6.0
216218_s_at	phospholipase C-like 2	PLCL2	23228	1.62E-06	4.16E-05	6.54	3.97	-6.0
210655_s_at fo	rkhead box O3	FOXO3	2309	1.62E-06	4.16E-05	7.91	5.33	-6.0
241929_at Tran	scribed locus	NA	NA	4.49E-06	5.50E-05	8.00	5.43	-6.0
1568648_a_at	CDNA clone IMAGE:4426835 NA		NA	4.49E-06	5.50E-05	6.36	3.78	-6.0
205350_at	cellular retinoic acid binding protein 1	CRABP1	1381	1.62E-06	4.16E-05	8.27	5.70	-5.9
201043_s_at	acidic (leucine-rich) nuclear phosphoprotein 32 family, member A	ANP32A 8	125	1.62E-06	4.16E-05	7.59	5.02	-5.9
207002_s_at	pleiomorphic adenoma gene-like 1	PLAGL1	5325	1.62E-06	4.16E-05	9.78	7.22	-5.9

238794_at	chromosome 10 open reading frame 78	C10orf78	119392	1.62E-06	4.16E-05	6.77	4.21	-5.9
222719_s_at	platelet derived growth factor C	PDGFC	56034	1.62E-06	4.16E-05	7.79	5.24	-5.9
208867_s_at	casein kinase 1, alpha 1	CSNK1A1	1452	1.62E-06	4.16E-05	9.44	6.88	-5.9
211965_at	zinc finger protein 36, C3H type-like 1	ZFP36L1	677	1.62E-06	4.16E-05	8.19	5.64	-5.9
228885_at	MAM domain containing 2	MAMDC2	256691	1.93E-06	4.16E-05	8.87	6.32	-5.9
228703_at	procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), alpha polypeptide III	P4HA3	283208	5.30E-06	6.05E-05	6.06	8.61	5.9
218036_x_at	NMD3 homolog (S. cerevisiae) NMD3		51068	1.62E-06	4.16E-05	8.52	5.98	-5.8
210764_s_at	cysteine-rich, angiogenic inducer, 61	CYR61	3491	6.25E-06	6.66E-05	11.01	8.47	-5.8
211022_s_at	alpha thalassemia/mental retardation syndrome X-linked (RAD54 homolog, S. cerevisiae)	ATRX 54	6	1.62E-06	4.16E-05	5.96	3.42	-5.8
233314_at	phosphatase and tensin homolog (mutated in multiple advanced cancers 1)	PTEN	5728	1.62E-06	4.16E-05	7.50	4.97	-5.8
208079_s_at au	rora kinase A	AURKA	6790	1.62E-06	4.16E-05	5.89	8.43	5.8
205666_at	flavin containing monooxygenase 1	FMO1	2326	4.49E-06	5.50E-05	8.63	6.11	-5.8
217546_at met	allothionein 1M	MT1M	4499	3.80E-06	5.03E-05	7.09	4.56	-5.8
216100_s_at	torsin A interacting protein 1	TOR1AIP1	26092	1.62E-06	4.16E-05	7.21	4.68	-5.8
230597_at	solute carrier family 7 (cationic amino acid transporter, y+ system), member 3	SLC7A3	84889	1.62E-06	4.16E-05	7.36	4.84	-5.7
204955_at	sushi-repeat-containing protein, X-linked	SRPX 8	406	1.62E-06	4.16E-05	11.26	8.74	-5.7
220266_s_at	Kruppel-like factor 4 (gut)	KLF4	9314	1.62E-06	4.16E-05	7.59	5.07	-5.7
218711_s_at	serum deprivation response (phosphatidylserine binding protein)	SDPR 8	436	1.62E-06	4.16E-05	7.36	4.84	-5.7
235616_at	teashirt zinc finger homeobox 2	TSHZ2	128553	1.62E-06	4.16E-05	9.47	6.96	-5.7
1554167_a_at	golgi autoantigen, golgin subfamily a, 7	GOLGA7	51125	1.62E-06	4.16E-05	9.20	6.69	-5.7
205518_s_at	cytidine monophosphate-N-acetylneuraminic acid hydroxylase (CMP-N-acetylneuraminate monooxygenase) CMA	H	8418	1.62E-06	4.16E-05	8.25	5.74	-5.7
209754_s_at t	hymopoietin	TMPO	7112	1.62E-06	4.16E-05	7.19	4.69	-5.7
202731_at	programmed cell death 4 (neoplastic transformation inhibitor)	PDCD4 27	250	7.36E-06	7.39E-05	10.53	8.02	-5.7
205479_s_at	plasminogen activator, urokinase PLAU		5328	1.62E-06	4.16E-05	7.65	10.14	5.7
1558136_s_at	TAF11 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 28kDa T	AF11	6882	1.62E-06	4.16E-05	8.28	5.79	-5.6
239218_at	CDNA FLJ43039 fis, clone BRTHA3003023 NA		NA	1.62E-06	4.16E-05	6.38	3.88	-5.6
1554470_s_at	zinc finger and BTB domain containing 44	ZBTB44	29068	1.62E-06	4.16E-05	7.52	5.03	-5.6
237860_at	Full-length cDNA clone CS0DC006YD17 of Neuroblastoma Cot 25-normalized of Homo sapiens (human)	NA NA		1.62E-06	4.16E-05	6.84	4.35	-5.6
204731_at	transforming growth factor, beta receptor III	TGFBR3	7049	1.62E-06	4.16E-05	10.74	8.26	-5.6
1552660_a_at	chromosome 5 open reading frame 22	C5orf22	55322	1.62E-06	4.16E-05	7.15	4.67	-5.6
216591_s_at	hCG1776980 | succinate dehydrogenase complex, subunit C, integral membrane protein, 15kDa	SDHC | hCG_1776980	6391 | 642502 1	.62E-06	4.16E-05	8.50	6.01	-5.6

228956_at	UDP glycosyltransferase 8 (UDP-galactose ceramide galactosyltransferase) UGT8		7368	1.93E-06	4.16E-05	5.58	3.10	-5.6
231955_s_at	3-hydroxyisobutyrate dehydrogenase	HIBADH 1	1112	1.62E-06	4.16E-05	8.11	5.63	-5.6
205200_at	C-type lectin domain family 3, member B	CLEC3B	7123	1.62E-06	4.16E-05	9.00	6.53	-5.6
238577_s_at	Teashirt zinc finger homeobox 2	TSHZ2	128553	1.62E-06	4.16E-05	8.99	6.52	-5.6
1556364_at	Hypothetical protein LOC730057	LOC730057 7	30057	1.62E-06	4.16E-05	7.92	5.44	-5.6
216997_x_at	transducin-like enhancer of split 4 (E(sp1) homolog, Drosophila)	TLE4 7	091	1.62E-06	4.16E-05	6.00	3.53	-5.6
238469_at	MRNA full length insert cDNA clone EUROIMAGE 1509279	NA NA		3.21E-06	4.66E-05	7.13	4.66	-5.6
226407_at	Pp13759 | Similar to Reticulocalbin-1 precursor	LOC728913	728913	1.62E-06	4.16E-05	7.28	4.81	-5.5
1553148_a_at	sorting nexin 13	SNX13	23161	1.62E-06	4.16E-05	7.28	4.82	-5.5
204753_s_at h	epatic leukemia factor	HLF	3131	3.80E-06	5.03E-05	7.72	5.26	-5.5
202920_at	ankyrin 2, neuronal	ANK2	287	2.71E-06	4.39E-05	8.38	5.91	-5.5
224046_s_at p	hosphodiesterase 7A	PDE7A	5150	1.62E-06	4.16E-05	6.81	4.35	-5.5
205287_s_at	transcription factor AP-2 gamma (activating enhancer binding protein 2 gamma)	TFAP2C 7	022	1.62E-06	4.16E-05	7.99	5.53	-5.5
209131_s_at	synaptosomal-associated protein, 23kDa	SNAP23 8	773	1.62E-06	4.16E-05	7.43	4.97	-5.5
203953_s_at cl	audin 3	CLDN3	1365	3.21E-06	4.66E-05	8.49	6.04	-5.5
239432_at	hypothetical protein FLJ31306 FLJ	31306	379025	1.62E-06	4.16E-05	8.66	6.21	-5.5
234491_s_at	salvador homolog 1 (Drosophila) SAV1		60485	1.62E-06	4.16E-05	9.69	7.23	-5.5
202516_s_at	discs, large homolog 1 (Drosophila) DLG1		1739	1.62E-06	4.16E-05	6.96	4.51	-5.5
224582_s_at	nuclear casein kinase and cyclin-dependent kinase substrate 1	NUCKS1 64	710	1.62E-06	4.16E-05	9.87	7.42	-5.5
200787_s_at	phosphoprotein enriched in astrocytes 15	PEA15	8682	1.62E-06	4.16E-05	8.41	5.96	-5.5
224339_s_at ang	iopoietin-like 1	ANGPTL1	9068	3.21E-06	4.66E-05	7.59	5.14	-5.4
1565269_s_at	activating transcription factor 1	ATF1	466	1.62E-06	4.16E-05	7.72	5.27	-5.4
1554767_s_at	crystallin, zeta (quinone reductase)-like 1	CRYZL1	9946	1.62E-06	4.16E-05	6.77	4.33	-5.4
210285_x_at	Wilms tumor 1 associated protein	WTAP 9	589	1.62E-06	4.16E-05	8.70	6.27	-5.4
240383_at	ubiquitin-conjugating enzyme E2D 3 (UBC4/5 homolog, yeast)	UBE2D3 7	323	1.62E-06	4.16E-05	8.65	6.22	-5.4
225381_at	hypothetical gene supported by BX647608	LOC399959	399959	1.62E-06	4.16E-05	9.66	7.23	-5.4
213183_s_at	Cyclin-dependent kinase inhibitor 1C (p57, Kip2)	CDKN1C 1	028	1.62E-06	4.16E-05	7.12	4.69	-5.4
229684_s_at	Zinc finger protein 644	ZNF644	84146	1.62E-06	4.16E-05	6.90	4.47	-5.4
204128_s_at	replication factor C (activator 1) 3, 38kDa	RFC3	5983	1.62E-06	4.16E-05	6.46	4.04	-5.4
209894_at l	eptin receptor	LEPR	3953	3.80E-06	5.03E-05	9.54	7.11	-5.4
237727_at	Full length insert cDNA YN61C04	NA NA		6.25E-06	6.66E-05	7.79	5.36	-5.4
211602_s_at	transient receptor potential cation channel, subfamily C, member 1	TRPC1 7	220	1.62E-06	4.16E-05	5.96	3.54	-5.4
1552344_s_at	CCR4-NOT transcription complex, subunit 7	CNOT7	29883	1.93E-06	4.16E-05	7.34	4.92	-5.4
226094_at	phosphoinositide-3-kinase, class 2, alpha polypeptide	PIK3C2A	5286	1.62E-06	4.16E-05	8.57	6.15	-5.3
209169_at g	lycoprotein M6B	GPM6B	2824	1.62E-06	4.16E-05	8.96	6.54	-5.3
225016_at	adenomatosis polyposis coli down-regulated 1	APCDD1	147495	1.62E-06	4.16E-05	8.99	6.57	-5.3
205746_s_at	ADAM metallopeptidase domain 17 (tumor necrosis	ADAM17	6868	1.62E-06	4.16E-05	7.24	4.82	-5.3

	factor, alpha, converting enzyme)							
229110_at	CDNA clone IMAGE:4794876 NA		NA	1.62E-06	4.16E-05	3.55	5.97	5.3
1552685_a_at	grainyhead-like 1 (Drosophila) GRHL1		29841	1.62E-06	4.16E-05	7.29	4.87	-5.3
205051_s_at	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	KIT 3	815	1.92E-06	4.16E-05	10.53	8.13	-5.3
211536_x_at	mitogen-activated protein kinase kinase kinase 7	MAP3K7 6	885	1.62E-06	4.16E-05	7.69	5.28	-5.3
228033_at	E2F transcription factor 7	E2F7	144455	2.71E-06	4.39E-05	4.19	6.60	5.3
238049_at	GRAM domain containing 3	GRAMD3	65983	5.30E-06	6.05E-05	8.29	5.88	-5.3
228584_at	sarcoglycan, beta (43kDa dystrophin-associated glycoprotein)	SGCB 6	443	1.62E-06	4.16E-05	7.68	5.28	-5.3
221194_s_at PTD0	16 protein	LOC51136	51136	1.62E-06	4.16E-05	6.73	4.33	-5.3
214091_s_at	glutathione peroxidase 3 (plasma) GPX3		2878	5.30E-06	6.05E-05	8.77	6.37	-5.3
219492_at	cysteine-rich hydrophobic domain 2	CHIC2	26511	1.62E-06	4.16E-05	9.40	7.00	-5.3
210881_s_at	insulin- insulin-like growth factor 2 | insulin-like growth factor 2 (somatomedin A)	IGF2 | INS-IGF2	3481 | 723961	8.65E-06 8	.19E-05	7.29	4.89	-5.3
217197_x_at hy	pothetical gene CG018	CG018	90634	1.62E-06	4.16E-05	7.21	4.81	-5.3
203881_s_at	dystrophin (muscular dystrophy, Duchenne and Becker types)	DMD	1756	1.93E-06	4.16E-05	8.79	6.39	-5.3
243041_s_at Tran	scribed locus	NA	NA	2.29E-06	4.19E-05	8.52	6.13	-5.3
215787_at	Actin, alpha 2, smooth muscle, aorta	ACTA2	59	1.62E-06	4.16E-05	5.74	3.35	-5.3
232204_at	early B-cell factor 1	EBF1	1879	2.29E-06	4.19E-05	8.30	5.91	-5.3
207943_x_at	pleiomorphic adenoma gene-like 1	PLAGL1 5	325	1.62E-06	4.16E-05	9.12	6.73	-5.2
1559949_at	Full length insert cDNA clone YA84A05	NA	NA	1.62E-06	4.16E-05	6.09	8.48	5.2
211749_s_at	vesicle-associated membrane protein 3 (cellubrevin)	VAMP3	9341	1.62E-06	4.16E-05	9.88	7.49	-5.2
228416_at	activin A receptor, type IIA	ACVR2A	92	1.62E-06	4.16E-05	8.79	6.40	-5.2
229691_at	Full-length cDNA clone CS0DH005YI18 of T cells (Jurkat cell line) of Homo sapiens (human)	NA	NA	1.62E-06	4.16E-05	6.44	8.83	5.2
238178_at Tran	scribed locus	NA	NA	1.93E-06	4.16E-05	5.90	3.51	-5.2
241726_at Tran	scribed locus	NA	NA	4.49E-06	5.50E-05	4.20	6.59	5.2
215990_s_at	B-cell CLL/lymphoma 6 (zinc finger protein 51)	BCL6	604	1.62E-06	4.16E-05	7.02	4.64	-5.2
211896_s_at d	ecorin	DCN	1634	1.62E-06	4.16E-05	12.86	10.48	-5.2
203065_s_at	caveolin 1, caveolae protein, 22kDa CA	V1	857	1.62E-06	4.16E-05	10.99	8.61	-5.2
200730_s_at	protein tyrosine phosphatase type IVA, member 1	PTP4A1	7803	2.29E-06	4.19E-05	8.67	6.29	-5.2
212713_at	microfibrillar-associated protein 4	MFAP4	4239	1.93E-06	4.16E-05	9.03	6.65	-5.2
216915_s_at	protein tyrosine phosphatase, non-receptor type 12	PTPN12	5782	1.93E-06	4.16E-05	6.69	4.31	-5.2
203697_at fri	zzled-related protein	FRZB	2487	1.62E-06	4.16E-05	9.66	7.28	-5.2
235355_at	MRNA; cDNA DKFZp564E143 (from clone DKFZp564E143)	NA NA		2.71E-06	4.39E-05	8.18	5.81	-5.2
228750_at Tran	scribed locus	NA	NA	6.25E-06	6.66E-05	10.24	7.86	-5.2
216048_s_at	Rho-related BTB domain containing 3	RHOBTB3	22836	1.62E-06	4.16E-05	7.04	4.67	-5.2
210571_s_at	cytidine monophosphate-N-acetylneuraminic acid hydroxylase (CMP-N-acetylneuraminate monooxygenase) CMA	H	8418	1.62E-06	4.16E-05	6.72	4.34	-5.2
209209_s_at	pleckstrin homology domain containing, family C (with FERM domain) member 1	PLEKHC1 10	979	1.62E-06	4.16E-05	8.23	5.86	-5.2

205619_s_at m	esenchyme homeobox 1	MEOX1	4222	1.02E-05	9.16E-05	7.75	5.38	-5.2
222038_s_at	Non-metastatic cells 1, protein (NM23A) expressed in	NME1 4	830	1.62E-06	4.16E-05	6.03	3.66	-5.2
226834_at Tran	scribed locus	NA	NA	1.62E-06	4.16E-05	9.27	6.90	-5.2
219497_s_at	B-cell CLL/lymphoma 11A (zinc finger protein)	BCL11A	53335	6.25E-06	6.66E-05	8.13	5.76	-5.2
1569433_at	sterile alpha motif domain containing 5	SAMD5	389432	1.62E-06	4.16E-05	5.65	3.28	-5.2
1553530_a_at	integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29 includes MDF2, MSK12)	ITGB1	3688	1.62E-06	4.16E-05	10.93	8.57	-5.2
242626_at	sterile alpha motif domain containing 5	SAMD5 3	89432	6.25E-06	6.66E-05	8.25	5.89	-5.2
203744_at	high-mobility group box 3	HMGB3	3149	1.62E-06	4.16E-05	7.43	9.80	5.2
1555812_a_at	Rho GDP dissociation inhibitor (GDI) beta	ARHGDIB	397	1.62E-06	4.16E-05	9.30	6.93	-5.1
234492_at	CDNA FLJ11818 fis, clone HEMBA1006424	NA NA		1.62E-06	4.16E-05	5.18	2.81	-5.1
241774_at Tran	scribed locus	NA	NA	1.62E-06	4.16E-05	8.12	5.76	-5.1
212187_x_at	prostaglandin D2 synthase 21kDa (brain)	PTGDS	5730	1.02E-05	9.16E-05	9.78	7.42	-5.1
213022_s_at u	trophin	UTRN	7402	1.62E-06	4.16E-05	7.78	5.42	-5.1
230000_at	ring finger protein 213	RNF213	57674	1.93E-06	4.16E-05	6.05	8.40	5.1
219773_at	NADPH oxidase 4	NOX4	50507	8.65E-06	8.19E-05	4.85	7.20	5.1
222611_s_at para	speckle component 1	PSPC1	55269	1.62E-06	4.16E-05	6.33	3.98	-5.1
228634_s_at	Cold shock domain protein A CSDA		8531	1.62E-06	4.16E-05	6.55	4.20	-5.1
210756_s_at	Notch homolog 2 (Drosophila) NOTCH2		4853	1.62E-06	4.16E-05	8.89	6.55	-5.1
210786_s_at	Friend leukemia virus integration 1	FLI1	2313	1.62E-06	4.16E-05	6.76	4.41	-5.1
244581_at	Zinc finger and BTB domain containing 20	ZBTB20	26137	1.62E-06	4.16E-05	7.44	5.09	-5.1
225990_at Boc	homolog (mouse)	BOC	91653	1.62E-06	4.16E-05	10.17	7.82	-5.1
210338_s_at	heat shock 70kDa protein 8 | similar to heat shock protein 8	HSPA8 | LOC402143	3312 | 402143 1	.62E-06	4.16E-05	11.94	9.60	-5.1
212595_s_at	DAZ associated protein 2	DAZAP2	9802	1.62E-06	4.16E-05	10.46	8.11	-5.1
215891_s_at	GM2 ganglioside activator	GM2A	2760	1.62E-06	4.16E-05	6.60	4.25	-5.1
223661_at NA		NA	NA	1.62E-06	4.16E-05	6.36	4.01	-5.1
1568647_at	CDNA clone IMAGE:4426835 NA		NA	5.30E-06	6.05E-05	6.28	3.94	-5.1
239155_at	similar to coxsackie virus and adenovirus receptor precursor	LOC653108 | LOC730425	653108 | 730425	2.71E-06 4	.39E-05	5.62	7.96	5.1
200769_s_at	methionine adenosyltransferase II, alpha	MAT2A	4144	1.62E-06	4.16E-05	9.06	6.72	-5.1
1569106_s_at	SET domain containing 5	SETD5	55209	1.93E-06	4.16E-05	6.34	4.00	-5.1
214691_x_at	family with sequence similarity 63, member B	FAM63B	54629	1.62E-06	4.16E-05	6.82	4.48	-5.1
209055_s_at	CDC5 cell division cycle 5-like (S. pombe)	CDC5L	988	1.62E-06	4.16E-05	7.89	5.56	-5.1
205612_at m	ultimerin 1	MMRN1	22915	2.28E-06	4.19E-05	5.30	2.96	-5.0
244433_at NA		NA	NA	1.62E-06	4.16E-05	8.34	6.01	-5.0
217728_at	S100 calcium binding protein A6	S100A6	6277	2.29E-06	4.19E-05	11.50	9.16	-5.0
201149_s_at	TIMP metallopeptidase inhibitor 3 (Sorsby fundus dystrophy, pseudoinflammatory) TIMP3		7078	1.62E-06	4.16E-05	10.96	8.63	-5.0
1563466_at	myosin, light chain kinase	MYLK	4638	1.62E-06	4.16E-05	6.23	3.90	-5.0
231989_s_at	PI-3-kinase-related kinase SMG-1 - like locus | hypothetical protein LOC440345 | similar to PI-3-kinase-related kinase	LOC440345 | LOC641298 | LOC730099	440345 | 641298 | 730099 1	.62E-06	4.16E-05	9.28	6.95	-5.0

	SMG-1							
226825_s_at	transmembrane protein 165	TMEM165	55858	1.62E-06	4.16E-05	10.01	7.69	-5.0
233878_s_at	5'-3' exoribonuclease 2	XRN2	22803	1.62E-06	4.16E-05	8.31	5.99	-5.0
209505_at	Nuclear receptor subfamily 2, group F, member 1	NR2F1	7025	2.29E-06	4.19E-05	10.18	7.86	-5.0
229308_at Tran	scribed locus	NA	NA	1.62E-06	4.16E-05	6.93	4.61	-5.0
239301_at h	ypothetical LOC644285	LOC644285	644285	1.62E-06	4.16E-05	6.13	3.82	-5.0
1555996_s_at	eukaryotic translation initiation factor 4A, isoform 2 EIF4	A2	1974	1.62E-06	4.16E-05	7.54	5.22	-5.0
211748_x_at	prostaglandin D2 synthase 21kDa (brain)	PTGDS	5730	1.02E-05	9.16E-05	10.41	8.09	-5.0
217465_at	NCK-associated protein 1	NCKAP1	10787	1.62E-06	4.16E-05	7.10	4.78	-5.0
207781_s_at	zinc finger protein 711	ZNF711	7552 1	.62E-06	4.16E-05	5.55	3.24	-5.0
1567032_s_at	zinc finger protein 160	ZNF160	90338 1	.62E-06	4.16E-05	5.43	3.11	-5.0
227646_at	CDNA FLJ39389 fis, clone PLACE6003621	NA NA		3.21E-06	4.66E-05	10.24	7.93	-5.0
218236_s_at	protein kinase D3	PRKD3	23683	1.62E-06	4.16E-05	9.38	7.07	-5.0
1554574_a_at	cytochrome b5 reductase 3	CYB5R3	1727	1.62E-06	4.16E-05	8.14	5.83	-5.0
205961_s_at	PC4 and SFRS1 interacting protein 1	PSIP1	11168	1.62E-06	4.16E-05	9.28	6.97	-5.0
207782_s_at	presenilin 1 (Alzheimer disease 3)	PSEN1 5	663	1.62E-06	4.16E-05	7.41	5.10	-5.0
203435_s_at	membrane metallo-endopeptidase MME		4311	2.71E-06	4.39E-05	8.52	6.22	-4.9
225977_at p	rotocadherin 18	PCDH18	54510	1.62E-06	4.16E-05	7.70	5.40	-4.9
1555411_a_at cy	clin L1	CCNL1	57018	1.62E-06	4.16E-05	10.08	7.77	-4.9
227705_at	transcription elongation factor A (SII)-like 7	TCEAL7	56849	3.21E-06	4.66 E-05	9.15	6.85	-4.9
1552978_a_at	secretory carrier membrane protein 1	SCAMP1	9522	1.62E-06	4.16E-05	8.59	6.29	-4.9
212354_at su	lfatase 1	SULF1	23213	1.62E-06	4.16E-05	8.47	10.76	4.9
1555193_a_at	zinc finger protein 277 pseudogene	ZNF277P 1	1179	1.62E-06	4.16E-05	6.40	4.10	-4.9
204438_at	mannose receptor, C type 1 | mannose receptor, C type 1-like 1	MRC1 | MRC1L1	414308 | 4360	2.29E-06	4 .19E-05	8.77	6.47	-4.9
244042_x_at	Similar to retinoic acid receptor responder (tazarotene induced) 2	LOC651466 6	51466	5.30E-06	6.05E-05	6.15	3.85	-4.9
1555408_at	B melanoma antigen family, member 2 | B melanoma antigen family, member 4	BAGE2 | BAGE4	85317 | 85319	1.62E-06	4 .16E-05	6.12	3.83	-4.9
233555_s_at s	ulfatase 2	SULF2	55959	1.62E-06	4.16E-05	7.52	5.23	-4.9
240773_at Tran	scribed locus	NA	NA	1.62E-06	4.16E-05	3.16	5.45	4.9
214543_x_at	quaking homolog, KH domain RNA binding (mouse)	QKI 9	444	1.62E-06	4.16E-05	9.17	6.88	-4.9
212272_at li	pin 1	LPIN1	23175	1.62E-06	4.16E-05	5.86	3.58	-4.9
236277_at	Primary neuroblastoma cDNA, clone:Nbla04246, full insert sequence	NA NA		1.62E-06	4.16E-05	6.18	3.89	-4.9
223620_at	G protein-coupled receptor 34 G	PR34	2857	1.62E-06	4.16E-05	8.87	6.59	-4.9
211356_x_at l	eptin receptor	LEPR	3953	1.62E-06	4.16E-05	6.93	4.65	-4.9
238474_at n	ucleoporin 43kDa	NUP43	348995	1.62E-06	4.16E-05	6.18	3.90	-4.9
206560_s_at	melanoma inhibitory activity	MIA	8190	3.21E-06	4.66E-05	8.77	6.49	-4.8
209948_at	potassium large conductance calcium-activated channel, subfamily M, beta member 1	KCNMB1 3	779	1.62E-06	4.16E-05	8.27	5.99	-4.8
205882_x_at	adducin 3 (gamma)	ADD3	120	1.62E-06	4.16E-05	11.05	8.77	-4.8
204369_at	phosphoinositide-3-kinase, catalytic, alpha polypeptide	PIK3CA 5	290	1.93E-06	4.16E-05	8.55	6.28	-4.8
219552_at	sushi, von Willebrand factor type A, EGF and pentraxin domain containing 1	SVEP1 79	987	1.62E-06	4.16E-05	5.71	3.43	-4.8

237746_at	Splicing factor, arginine/serine-rich 11	SFRS11	9295	1.62E-06	4.16E-05	5.87	3.60	-4.8
234645_at	CDNA: FLJ21664 fis, clone COL08890 NA		NA	1.62E-06	4.16E-05	6.53	4.25	-4.8
1555202_a_at	hypothetical protein FLJ10656 P1	5RS	55197	1.62E-06	4.16E-05	6.41	4.14	-4.8
208062_s_at n	euregulin 2	NRG2	9542	2.29E-06	4.19E-05	5.79	3.52	-4.8
201752_s_at	adducin 3 (gamma)	ADD3	120	1.62E-06	4.16E-05	11.06	8.80	-4.8
237411_at	ADAM metallopeptidase with thrombospondin type 1 motif, 6	ADAMTS6	11174	3.21E-06	4.66E-05	4.46	6.72	4.8
239262_at	CDNA FLJ26242 fis, clone DMC00770 NA		NA	5.30E-06	6.05E-05	8.04	5.77	-4.8
202017_at	epoxide hydrolase 1, microsomal (xenobiotic)	EPHX1	2052	1.93E-06	4.16E-05	7.71	5.45	-4.8
243648_at NA		NA	NA	1.62E-06	4.16E-05	7.27	9.54	4.8
1552721_a_at	fibroblast growth factor 1 (acidic) FGF1		2246	2.29E-06	4.19E-05	6.77	4.51	-4.8
1553672_at	enabled homolog (Drosophila) ENAH		55740	1.62E-06	4.16E-05	4.96	7.22	4.8
230670_at	immunoglobulin superfamily, member 10	IGSF10 2	85313	2.71E-06	4.39E-05	8.15	5.89	-4.8
216266_s_at	ADP-ribosylation factor guanine nucleotide-exchange factor 1(brefeldin A-inhibited) ARFGEF1		10565	1.62E-06	4.16E-05	7.90	5.64	-4.8
225664_at	collagen, type XII, alpha 1	COL12A1	1303	1.62E-06	4.16E-05	9.85	12.11	4.8
207016_s_at	aldehyde dehydrogenase 1 family, member A2	ALDH1A2	8854	7.36E-06	7.39E-05	7.49	5.24	-4.8
201340_s_at	ectodermal-neural cortex (with BTB-like domain)	ENC1	8507	3.21E-06	4.66E-05	6.10	8.36	4.8
203851_at	insulin-like growth factor binding protein 6	IGFBP6	3489	1.62E-06	4.16E-05	9.78	7.53	-4.8
205475_at	scrapie responsive protein 1	SCRG1	11341	1.62E-06	4.16E-05	6.87	4.61	-4.8
219890_at	C-type lectin domain family 5, member A	CLEC5A	23601	2.29E-06	4.19E-05	3.23	5.49	4.8
215294_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 1	SMARCA1	6594	5.30E-06	6.05E-05	8.12	5.87	-4.8
235446_at	X (inactive)-specific transcript, antisense	TSIX	9383	1.62E-06	4.16E-05	4.73	6.97	4.7
219534_x_at	cyclin-dependent kinase inhibitor 1C (p57, Kip2)	CDKN1C	1028	1.62E-06	4.16E-05	8.16	5.92	-4.7
210734_x_at	MYC associated factor X	MAX	4149	1.62E-06	4.16E-05	8.07	5.82	-4.7
210995_s_at	tripartite motif-containing 23	TRIM23	373	1.62E-06	4.16E-05	6.87	4.62	-4.7
1554676_at s	erglycin	SRGN	5552	4.49E-06	5.50E-05	6.87	4.62	-4.7
212099_at	ras homolog gene family, member B	RHOB	388	1.62E-06	4.16E-05	10.73	8.49	-4.7
230538_at	SHC (Src homology 2 domain containing) family, member 4	SHC4 3	99694	3.80E-06	5.03E-05	6.80	4.56	-4.7
1570021_at h	omeobox C14	LOC360030	360030	1.62E-06	4.16E-05	3.29	5.53	4.7
1554397_s_at	UEV and lactate/malate dehyrogenase domains	UEVLD	55293	1.62E-06	4.16E-05	7.52	5.28	-4.7
1558828_s_at	Hypothetical protein DKFZp586C0721	DKFZp586C072 1 1	53688	2.71E-06	4.39E-05	8.77	6.54	-4.7
221311_x_at	LYR motif containing 2	LYRM2	57226	1.62E-06	4.16E-05	8.62	6.38	-4.7
1553678_a_at	integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29 includes MDF2, MSK12)	ITGB1	3688	1.62E-06	4.16E-05	11.18	8.95	-4.7
239511_s_at	splicing factor, arginine/serine-rich 4	SFRS4 6	429	1.62E-06	4.16E-05	5.91	3.68	-4.7
223925_s_at	myeloproliferative disease-associated SEREX antigen	LOC767558	767558	1.62E-06	4.16E-05	6.45	4.22	-4.7
76897_s_at	FK506 binding protein 15, 133kDa FKBP15		23307	1.93E-06	4.16E-05	6.02	3.79	-4.7
204344_s_at	Sec23 homolog A (S. cerevisiae) SEC2	3A	10484	1.62E-06	4.16E-05	7.80	5.57	-4.7
1555978_s_at	CDNA FLJ33153 fis, clone UTERU2000332 NA		NA	5.30E-06	6.05E-05	6.33	4.10	-4.7

223585_x_at	kelch repeat and BTB (POZ) domain containing 2	KBTBD2	25948	1.62E-06	4.16E-05	7.68	5.45	-4.7
1564053_a_at	YTH domain family, member 3	YTHDF3	253943	1.62E-06	4.16E-05	7.28	5.06	-4.7
229620_at	selenoprotein P, plasma, 1	SEPP1	6414	1.93E-06	4.16E-05	7.15	4.93	-4.7
202075_s_at	phospholipid transfer protein	PLTP	5360	1.93E-06	4.16E-05	8.55	6.33	-4.7
227148_at	pleckstrin homology domain containing, family H (with MyTH4 domain) member 2	PLEKHH2 1	30271	1.02E-05	9.16E-05	9.15	6.93	-4.6
1555247_a_at	Rap guanine nucleotide exchange factor (GEF) 6	RAPGEF6 51	735	2.29E-06	4.19E-05	7.97	5.76	-4.6
207993_s_at	calcium binding protein P22	CHP	11261	1.62E-06	4.16E-05	6.91	4.70	-4.6
208097_s_at	thioredoxin domain containing 1	TXNDC1	81542	1.62E-06	4.16E-05	9.25	7.04	-4.6
208151_x_at	DEAD (Asp-Glu-Ala-Asp) box polypeptide 17	DDX17	10521	1.62E-06	4.16E-05	8.20	5.99	-4.6
1554417_s_at	anterior pharynx defective 1 homolog A (C. elegans)	APH1A	51107	1.62E-06	4.16E-05	8.34	6.13	-4.6
204271_s_at	endothelin receptor type B	EDNRB	1910	2.71E-06	4.39E-05	9.71	7.50	-4.6
231640_at	LYR motif containing 5	LYRM5	144363	1.62E-06	4.16E-05	7.21	5.01	-4.6
227771_at	leukemia inhibitory factor receptor alpha	LIFR 3	977	1.62E-06	4.16E-05	8.43	6.23	-4.6
219777_at	GTPase, IMAP family member 6	GIMAP6	474344	2.29E-06	4.19E-05	8.96	6.75	-4.6
228062_at	nucleosome assembly protein 1-like 5	NAP1L5	266812	5.30E-06	6.05E-05	8.20	5.99	-4.6
213182_x_at	cyclin-dependent kinase inhibitor 1C (p57, Kip2)	CDKN1C 1	028	1.62E-06	4.16E-05	7.96	5.76	-4.6
235716_at Tran	scribed locus	NA	NA	1.62E-06	4.16E-05	10.04	7.83	-4.6
209293_x_at	inhibitor of DNA binding 4, dominant negative helix-loop-helix protein	ID4 3	400	2.71E-06	4.39E-05	8.45	6.25	-4.6
211355_x_at l	eptin receptor	LEPR	3953	1.62E-06	4.16E-05	6.55	4.35	-4.6
225911_at n	ephronectin	NPNT	255743	7.36E-06	7.39E-05	7.49	9.69	4.6
243481_at	ras homolog gene family, member J	RHOJ 5	7381	1.62E-06	4.16E-05	6.56	4.36	-4.6
229118_at	proline rich Gla (G-carboxyglutamic acid) 3 (transmembrane)	PRRG3 79	057	1.62E-06	4.16E-05	6.30	4.10	-4.6
225070_at	nuclear undecaprenyl pyrophosphate synthase 1 homolog (S. cerevisiae)	NUS1 1	16150	1.62E-06	4.16E-05	8.44	6.24	-4.6
1555977_at	CDNA FLJ33153 fis, clone UTERU2000332 NA		NA	1.93E-06	4.16E-05	7.62	5.42	-4.6
212515_s_at	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, X-linked	DDX3X	1654	1.62E-06	4.16E-05	10.09	7.90	-4.6
222507_s_at	TMEM9 domain family, member B	TMEM9B	56674	1.62E-06	4.16E-05	9.01	6.83	-4.5
207079_s_at	mediator complex subunit 6	MED6	10001	1.62E-06	4.16E-05	8.33	6.15	-4.5
227239_at	family with sequence similarity 126, member A	FAM126A 8	4668	1.62E-06	4.16E-05	10.16	7.97	-4.5
228697_at	histidine triad nucleotide binding protein 3	HINT3	135114	1.93E-06	4.16E-05	9.04	6.86	-4.5
1565717_s_at	fusion (involved in t(12;16) in malignant liposarcoma)	FUS	2521	1.62E-06	4.16E-05	9.06	6.88	-4.5
241536_at	Full length insert cDNA YO61A08 NA		NA	1.62E-06	4.16E-05	6.33	4.15	-4.5
235744_at	PTC7 protein phosphatase homolog (S. cerevisiae)	PPTC7	160760	1.62E-06	4.16E-05	5.22	3.04	-4.5
201559_s_at	chloride intracellular channel 4	CLIC4	25932	1.62E-06	4.16E-05	8.13	5.96	-4.5
1562031_at	Janus kinase 2 (a protein tyrosine kinase)	JAK2	3717	1.62E-06	4.16E-05	6.71	4.53	-4.5
205778_at	kallikrein-related peptidase 7	KLK7	5650	5.30E-06	6.05E-05	7.41	5.23	-4.5
215933_s_at	hematopoietically expressed homeobox HHEX		3087	5.30E-06	6.05E-05	7.07	4.90	-4.5
235629_at	Transcribed locus, strongly similar to XP_516072.1 similar to fibronectin 1 isoform 2 preproprotein; cold-insoluble globulin; migration-stimulating factor	NA	NA	6.25E-06	6.66E-05	4.88	7.05	4.5

	[Pan troglodytes]							
1557961_s_at	EF-hand calcium binding protein 1	EFCBP1	64168	1.62E-06	4.16E-05	6.70	4.53	-4.5
225576_at	chromosome 6 open reading frame 72	C6orf72	116254	1.62E-06	4.16E-05	10.02	7.85	-4.5
210994_x_at	tripartite motif-containing 23	TRIM23	373	1.62E-06	4.16E-05	7.23	5.06	-4.5
204845_s_at	glutamyl aminopeptidase (aminopeptidase A)	ENPEP	2028	1.62E-06	4.16E-05	6.07	3.90	-4.5
235591_at	somatostatin receptor 1	SSTR1	6751	6.25E-06	6.66E-05	6.18	4.01	-4.5
224368_s_at	NDRG family member 3	NDRG3	57446	1.62E-06	4.16E-05	7.06	4.89	-4.5
203357_s_at calp	ain 7	CAPN7	23473	1.62E-06	4.16E-05	8.63	6.47	-4.5
237737_at	hypothetical LOC375010 | hypothetical LOC401131 | hypothetical LOC643166 | hypothetical LOC643579 | hypothetical protein LOC728295 | hypothetical protein LOC728364 | hypothetical protein LOC728384 | hypothetical protein LOC728759 | hypothetical protein LOC728783 | similar to Ankyrin repeat domain-containing protein 18A	LOC375010 | LOC401131 | LOC643166 | LOC643579 | LOC646340 | LOC728295 | LOC728364 | LOC728384 | LOC728759 | LOC728783	375010 | 401131 | 643166 | 643579 | 646340 | 728295 | 728364 | 728384 | 728759 | 728783 2	.71E-06	4.39E-05	6.81	4.65	-4.5
212473_s_at	microtubule associated monoxygenase, calponin and LIM domain containing 2	MICAL2	9645	2.29E-06	4.19E-05	9.03	11.20	4.5
213548_s_at	CDV3 homolog (mouse)	CDV3	55573	1.93E-06	4.16E-05	6.81	4.65	-4.5
1555618_s_at	SUMO1 activating enzyme subunit 1	SAE1	10055	1.62E-06	4.16E-05	8.32	6.16	-4.5
1556657_at	CDNA FLJ36459 fis, clone THYMU2014762 NA		NA	1.62E-06	4.16E-05	8.45	6.29	-4.5
1569872_a_at	Full-length cDNA clone CS0DF015YK23 of Fetal brain of Homo sapiens (human) | Hypothetical protein LOC650392	LOC650392	650392	1.62E-06	4.16E-05	7.31	5.15	-4.5
204825_at	maternal embryonic leucine zipper kinase	MELK	9833	1.02E-05	9.16E-05	6.17	8.33	4.5
206157_at	pentraxin-related gene, rapidly induced by IL-1 beta	PTX3	5806	2.29E-06	4.19E-05	6.47	4.31	-4.5
219310_at	chromosome 20 open reading frame 39	C20orf39	79953	2.71E-06	4.39E-05	4.18	6.33	4.5
34031_i_at	KRIT1, ankyrin repeat containing K	RIT1	889	1.62E-06	4.16E-05	8.80	6.64	-4.4
242774_at	spectrin repeat containing, nuclear envelope 2	SYNE2	23224	1.62E-06	4.16E-05	7.29	5.14	-4.4
227006_at	protein phosphatase 1, regulatory (inhibitor) subunit 14A	PPP1R14A 94274		3.21E-06	4.66E-05 7.55		5.40 -4.4	
224549_x_at NA		NA	NA	1.62E-06	4.16E-05	8.21	6.05	-4.4
238905_at	ras homolog gene family, member J	RHOJ	57381	1.62E-06	4.16E-05	8.22	6.07	-4.4
212464_s_at fi	bronectin 1	FN1	2335	1.93E-06	4.16E-05	10.88	13.03	4.4
226220_at	Methyltransferase like 9	METTL9	51108 1	.62E-06	4.16E-05	7.44	5.29	-4.4
200624_s_at	matrin 3 | similar to Matrin-3 (Nuclear scaffold protein P130/MAT3)	LOC727839 | MATR3	727839 | 9782 1	.62E-06	4.16E-05	11.19	9.05	-4.4
210935_s_at	WD repeat domain 1	WDR1	9948	1.62E-06	4.16E-05	6.89	4.75	-4.4
204132_s_at fo	rkhead box O3	FOXO3	2309	1.62E-06	4.16E-05	8.47	6.33	-4.4
205421_at	solute carrier family 22 (extraneuronal monoamine transporter), member 3	SLC22A3 6	581	4.49E-06	5.50E-05	5.32	3.18	-4.4
202566_s_at s	upervillin	SVIL	6840	1.62E-06	4.16E-05	8.47	6.33	-4.4
238951_at Tran	scribed locus	NA	NA	1.62E-06	4.16E-05	6.26	4.12	-4.4
220955_x_at	RAB23, member RAS oncogene family	RAB23 51	715	1.62E-06	4.16E-05	7.17	5.03	-4.4
205188_s_at	SMAD family member 5	SMAD5	4090	1.62E-06	4.16E-05	7.53	5.39	-4.4

201497_x_at	myosin, heavy chain 11, smooth muscle	MYH11	4629	2.71E-06	4.39E-05	10.45	8.32	-4.4
232883_at	CDNA FLJ11977 fis, clone HEMBB1001254 NA		NA	1.62E-06	4.16E-05	6.28	4.15	-4.4
203798_s_at v	isinin-like 1	VSNL1	7447	3.21E-06	4.66E-05	5.95	3.81	-4.4
213887_s_at	polymerase (RNA) II (DNA directed) polypeptide E, 25kDa	POLR2E 5	434	1.62E-06	4.16E-05	8.66	6.53	-4.4
230440_at	zinc finger protein 469	ZNF469	84627 1	62E-06	4.16E-05	6.90	9.03	4.4
220199_s_at	chromosome 1 open reading frame 80	C1orf80 64	853	1.62E-06	4.16E-05	9.26	7.13	-4.4
205158_at	ribonuclease, RNase A family, 4	RNASE4	6038	1.62E-06	4.16E-05	10.10	7.97	-4.4
234032_at PRO	550	NA	NA	1.62E-06	4.16E-05	8.79	6.67	-4.4
217208_s_at	discs, large homolog 1 (Drosophila) DLG1		1739	1.93E-06	4.16E-05	7.30	5.17	-4.4
210639_s_at	ATG5 autophagy related 5 homolog (S. cerevisiae)	ATG5	9474	1.62E-06	4.16E-05	7.81	5.69	-4.4
209936_at RNA	binding motif protein 5	RBM5	10181	1.62E-06	4.16E-05	7.46	5.35	-4.3
228881_at	presenilin associated, rhomboid-like PA	RL	55486	1.62E-06	4.16E-05	7.03	4.92	-4.3
1554757_a_at	inositol polyphosphate-5-phosphatase, 40kDa	INPP5A	3632	1.62E-06	4.16E-05	6.27	4.16	-4.3
225088_at	chromosome 16 open reading frame 63	C16orf63	123811	1.62E-06	4.16E-05	7.32	5.21	-4.3
1554606_at	coiled-coil domain containing 100	CCDC100	153241	1.62E-06	4.16E-05	6.02	3.91	-4.3
201929_s_at p	lakophilin 4	PKP4	8502	1.62E-06	4.16E-05	8.96	6.85	-4.3
204755_x_at h	epatic leukemia factor	HLF	3131	1.02E-05	9.16E-05	8.18	6.07	-4.3
203323_at cav	eolin 2	CAV2	858	4.49E-06	5.50E-05	9.86	7.76	-4.3
228728_at	hypothetical protein FLJ21986 FLJ	21986	79974	3.21E-06	4.66E-05	8.70	6.60	-4.3
227529_s_at	A kinase (PRKA) anchor protein (gravin) 12	AKAP12 9	590	6.25E-06	6.66E-05	7.36	5.26	-4.3
216442_x_at fi	bronectin 1	FN1	2335	1.62E-06	4.16E-05	11.33	13.44	4.3
210203_at	CCR4-NOT transcription complex, subunit 4	CNOT4	4850	1.62E-06	4.16E-05	7.11	5.01	-4.3
219935_at	ADAM metallopeptidase with thrombospondin type 1 motif, 5 (aggrecanase-2)	ADAMTS5 1	1096	5.30E-06	6.05E-05	10.35	8.25	-4.3
231721_at	junctional adhesion molecule 3	JAM3	83700	1.62E-06	4.16E-05	5.50	3.40	-4.3
208925_at	claudin domain containing 1	CLDND1	56650	1.62E-06	4.16E-05	9.37	7.27	-4.3
244171_at	muskelin 1, intracellular mediator containing kelch motifs	MKLN1 4	289	1.62E-06	4.16E-05	5.34	3.24	-4.3
204092_s_at au	rora kinase A	AURKA	6790	1.62E-06	4.16E-05	5.80	7.89	4.3
202113_s_at	sorting nexin 2	SNX2	6643	1.62E-06	4.16E-05	9.55	7.45	-4.3
204036_at	endothelial differentiation, lysophosphatidic acid G-protein-coupled receptor, 2	EDG2	1902	3.80E-06	5.03E-05	9.16	7.07	-4.3
217915_s_at	chromosome 15 open reading frame 15	C15orf15	51187	1.62E-06	4.16E-05	10.64	8.55	-4.3
236038_at Tran	scribed locus	NA	NA	2.29E-06	4.19E-05	7.05	4.96	-4.3
218013_x_at d	ynactin 4 (p62)	DCTN4	51164	1.62E-06	4.16E-05	8.32	6.23	-4.3
205655_at	Mdm4, transformed 3T3 cell double minute 4, p53 binding protein (mouse)	MDM4	4194	3.21E-06	4.66E-05	6.16	4.07	-4.3
230156_x_at	Chromodomain helicase DNA binding protein 2	CHD2	1106	2.29E-06	4.19E-05	6.61	4.53	-4.2
206306_at ry	anodine receptor 3	RYR3	6263	1.62E-06	4.16E-05	5.80	3.71	-4.2
235463_s_at	LAG1 homolog, ceramide synthase 6	LASS6	253782	6.25E-06	6.66E-05	6.44	8.52	4.2
228133_s_at	myosin, heavy chain 11, smooth muscle	MYH11	4629	1.62E-06	4.16E-05	6.95	4.87	-4.2
233292_s_at	MASK-4E-BP3 alternate reading frame gene | ankyrin repeat and KH domain containing 1	ANKHD1 | MASK-BP3	404734 | 54882 1	.62E-06	4.16E-05	7.37	5.29	-4.2

1560342_at	CDNA clone IMAGE:5275043 NA		NA	1.93E-06	4.16E-05	7.04	4.96	-4.2
1567706_at	Clone TUB2 Cri-du-chat region mRNA	NA	NA	5.30E-06	6.05E-05	4.74	2.66	-4.2
216831_s_at	runt-related transcription factor 1; translocated to, 1 (cyclin D-related)	RUNX1T1 8	62	1.62E-06	4.16E-05	7.13	5.05	-4.2
242706_s_at	mediator complex subunit 23	MED23	9439	1.62E-06	4.16E-05	7.88	5.80	-4.2
209687_at	chemokine (C-X-C motif) ligand 12 (stromal cell-derived factor 1)	CXCL12 6	387	2.71E-06	4.39E-05	12.08	10.00	-4.2
228913_at	Pp13759 | Similar to Reticulocalbin-1 precursor	LOC728913 7	28913	1.62E-06	4.16E-05	8.04	5.96	-4.2
223674_s_at	CDC42 small effector 1	CDC42SE1	56882 1	.62E-06	4.16E-05	6.72	4.65	-4.2
222449_at	transmembrane, prostate androgen induced RNA	TMEPAI	56937	4.49E-06	5.50E-05	8.87	10.94	4.2
206042_x_at	SNRPN upstream reading frame | small nuclear ribonucleoprotein polypeptide N	SNRPN | SNURF	6638 | 8926 1	.62E-06	4.16E-05	9.27	7.20	-4.2
214129_at	similar to phosphodiesterase 4D interacting protein isoform 2	LOC727942	727942	1.62E-06	4.16E-05	6.87	8.94	4.2
229435_at	GLIS family zinc finger 3	GLIS3	169792	3.80E-06	5.03E-05	5.01	7.09	4.2
1555594_a_at	muscleblind-like (Drosophila) MBNL1		4154	1.62E-06	4.16E-05	5.97	3.90	-4.2
1555326_a_at	ADAM metallopeptidase domain 9 (meltrin gamma)	ADAM9	8754	4.49E-06	5.50E-05	7.64	5.57	-4.2
213844_at h	omeobox A5	HOXA5	3202	4.49E-06	5.50E-05	9.13	7.06	-4.2
210495_x_at fi	bronectin 1	FN1	2335	1.62E-06	4.16E-05	11.31	13.38	4.2
218212_s_at	molybdenum cofactor synthesis 2	MOCS2	4338	1.93E-06	4.16E-05	7.74	5.68	-4.2
205408_at	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 10	MLLT10 8	028	1.62E-06	4.16E-05	7.19	5.12	-4.2
225294_s_at	trafficking protein particle complex 1	TRAPPC1	58485	2.71E-06	4.39E-05	7.02	4.95	-4.2
214906_x_at hy	pothetical gene CG018	CG018	90634	1.62E-06	4.16E-05	7.23	5.17	-4.2
202278_s_at	serine palmitoyltransferase, long chain base subunit 1	SPTLC1	10558	1.62E-06	4.16E-05	8.26	6.20	-4.2
1554509_a_at	chromosome 10 open reading frame 97	C10orf97	80013	1.62E-06	4.16E-05	7.26	5.20	-4.2
1552289_a_at	cartilage intermediate layer protein 2	CILP2 1	48113	4.49E-06	5.50E-05	5.07	7.13	4.2
239245_at Tran	scribed locus	NA	NA	1.62E-06	4.16E-05	4.99	2.93	-4.2
243329_at	Clone IMAGE:121662 mRNA sequence	NA	NA	1.62E-06	4.16E-05	6.80	8.86	4.2
210543_s_at	protein kinase, DNA-activated, catalytic polypeptide	PRKDC 5	591	1.62E-06	4.16E-05	8.11	6.05	-4.2
218340_s_at	ubiquitin-activating enzyme E1-like 2	UBE1L2	55236	1.62E-06	4.16E-05	7.94	5.88	-4.2
1559883_s_at	SAM domain and HD domain 1	SAMHD1	25939	1.93E-06	4.16E-05	6.39	4.33	-4.2
225034_at	hypothetical protein LOC286167 LOC2	86167	286167	1.93E-06	4.16E-05	6.49	8.54	4.2
1555226_s_at	chromosome 1 open reading frame 43	C1orf43	25912	1.62E-06	4.16E-05	10.02	7.97	-4.1
211776_s_at	erythrocyte membrane protein band 4.1-like 3	EPB41L3	23136	6.25E-06	6.66E-05	7.95	5.90	-4.1
1569203_at	chemokine (C-X-C motif) ligand 2	CXCL2	2920	3.21E-06	4.66E-05	6.48	4.43	-4.1
1559477_s_at Meis	homeobox 1	MEIS1	4211	1.62E-06	4.16E-05	7.03	4.99	-4.1
203293_s_at l	ectin, mannose-binding, 1	LMAN1	3998	1.62E-06	4.16E-05	7.17	5.13	-4.1
216242_x_at	DNA directed RNA polymerase II polypeptide J-related | RPB11b2 protein | polymerase (RNA) II (DNA directed) polypeptide J, 13.3kDa pseudogene	POLR2J2 | POLR2J3 | POLR2J4	246721 | 548644 | 84820 1	.62E-06	4.16E-05	9.03	6.99	-4.1
1555780_a_at	Ras homolog enriched in brain RHEB		6009	1.62E-06	4.16E-05	8.97	6.92	-4.1

219834_at	amyotrophic lateral sclerosis 2 (juvenile) chromosome region, candidate 8	ALS2CR8	79800	1.62E-06	4.16E-05	6.41	4.37	-4.1
203375_s_at t	ripeptidyl peptidase II	TPP2	7174	1.62E-06	4.16E-05	8.74	6.71	-4.1
220794_at	gremlin 2, cysteine knot superfamily, homolog (Xenopus laevis)	GREM2	64388	3.80E-06	5.03E-05	6.24	4.21	-4.1
241820_at	RAP1 interacting factor homolog (yeast)	RIF1	55183	1.62E-06	4.16E-05	7.31	5.27	-4.1
1552695_a_at	solute carrier family 2 (facilitated glucose transporter), member 13	SLC2A13 1	14134	1.93E-06	4.16E-05	3.50	5.53	4.1
222678_s_at	DCN1, defective in cullin neddylation 1, domain containing 1 (S. cerevisiae)	DCUN1D1 5	4165	1.62E-06	4.16E-05	6.99	4.96	-4.1
227242_s_at	early B-cell factor 3	EBF3	253738 2	.29E-06	4.19E-05	5.57	3.54	-4.1
1554547_at	family with sequence similarity 13, member C1	FAM13C1	220965	5.30E-06	6.05E-05	7.87	5.84	-4.1
213913_s_at K	IAA0984 protein	KIAA0984	23329	1.62E-06	4.16E-05	5.82	7.85	4.1
228950_s_at	G protein-coupled receptor 177 G	PR177	79971	4.49E-06	5.50E-05	8.90	6.87	-4.1
215945_s_at	tripartite motif-containing 2	TRIM2	23321	1.62E-06	4.16E-05	7.83	5.80	-4.1
203543_s_at	Kruppel-like factor 9	KLF9	687	1.62E-06	4.16E-05	8.26	6.23	-4.1
210815_s_at cal	citonin receptor-like	CALCRL	10203	2.71E-06	4.39E-05	6.65	4.62	-4.1
223875_s_at	enhancer of polycomb homolog 1 (Drosophila)	EPC1	80314	1.62E-06	4.16E-05	6.67	4.64	-4.1
204480_s_at	chromosome 9 open reading frame 16	C9orf16 79	095	7.36E-06	7.39E-05	7.96	5.93	-4.1
1559255_a_at	Full length insert cDNA clone YW26C09	NA	NA	2.29E-06	4.19E-05	5.99	3.97	-4.1
230109_at p	hosphodiesterase 7B	PDE7B	27115	8.65E-06	8.19E-05	8.36	6.33	-4.1
216081_at	laminin, alpha 4	LAMA4	3910	2.71E-06	4.39E-05	7.10	5.07	-4.1
201523_x_at	ubiquitin-conjugating enzyme E2N (UBC13 homolog, yeast)	UBE2N	7334	1.62E-06	4.16E-05	9.18	7.15	-4.1
204866_at	PHD finger protein 16	PHF16	9767	1.93E-06	4.16E-05	7.31	9.33	4.1
1557169_x_at	HLA complex group 11	HCG11	493812	1.62E-06	4.16E-05	7.16	5.14	-4.1
215758_x_at	zinc finger protein 93	ZNF93	81931 1	.62E-06	4.16E-05	7.30	5.28	-4.1
233230_s_at	SLAIN motif family, member 2	SLAIN2	57606	1.62E-06	4.16E-05	5.54	3.52	-4.1
211676_s_at interf	eron gamma receptor 1	IFNGR1	3459 1	.62E-06	4.16E-05 10.	15	8.13	-4.1
1555097_a_at	prostaglandin F receptor (FP) PTGFR		5737	2.29E-06	4.19E-05	5.34	3.32	-4.1
229302_at	transmembrane protein 178	TMEM178	130733	1.93E-06	4.16E-05	8.19	6.17	-4.1
232125_at	CDNA FLJ12166 fis, clone MAMMA1000616 NA		NA	1.62E-06	4.16E-05	8.08	6.06	-4.1
201309_x_at	chromosome 5 open reading frame 13	C5orf13	9315	1.93E-06	4.16E-05	7.96	5.94	-4.1
229525_at NA		NA	NA	1.62E-06	4.16E-05	6.81	4.79	-4.0
221425_s_at	iron-sulfur cluster assembly 1 homolog (S. cerevisiae)	ISCA1	81689	1.62E-06	4.16E-05	7.56	5.54	-4.0
238490_at KIAA2	026	KIAA2026	158358	1.62E-06	4.16E-05	7.10	5.08	-4.0
200888_s_at	ribosomal protein L23	RPL23	9349	2.71E-06	4.39E-05	12.47	10.46	-4.0
210149_s_at	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit d	ATP5H 1	0476	2.71E-06	4.39E-05	11.15	9.13	-4.0
1557167_at	HLA complex group 11	HCG11	493812	1.62E-06	4.16E-05	6.99	4.98	-4.0
211719_x_at fi	bronectin 1	FN1	2335	1.62E-06	4.16E-05	11.43	13.44	4.0
1554899_s_at	Fc fragment of IgE, high affinity I, receptor for; gamma polypeptide	FCER1G 2	207	8.65E-06	8.19E-05	7.92	5.90	-4.0
202444_s_at	ER lipid raft associated 1	ERLIN1 1	0613	1.62E-06	4.16E-05	6.87	4.86	-4.0
228648_at	leucine-rich alpha-2-glycoprotein 1	LRG1	116844	1.02E-05	9.16E-05	6.31	8.32	4.0
235395_at	SEC63 homolog (S. cerevisiae) SEC6	3	11231	1.62E-06	4.16E-05	5.00	2.99	-4.0

209355_s_at	phosphatidic acid phosphatase type 2B	PPAP2B	8613	1.62E-06	4.16E-05	9.89	7.88	-4.0
231766_s_at	collagen, type XII, alpha 1	COL12A1	1303	4.49E-06	5.50E-05	7.21	9.22	4.0
1554450_s_at	mesoderm induction early response 1, family member 3	MIER3	166968 1	.62E-06	4.16E-05	6.59	4.59	-4.0
227984_at	Hypothetical protein LOC650392 LOC6	50392	650392	2.71E-06	4.39E-05	6.05	4.05	-4.0
222450_at	transmembrane, prostate androgen induced RNA	TMEPAI 56	937	7.36E-06	7.39E-05	8.26	10.26	4.0
227277_at	CDNA FLJ41088 fis, clone ASTRO2002459 NA		NA	1.62E-06	4.16E-05	7.57	9.56	4.0
215318_at hy	pothetical gene CG012	CG012	116829	1.62E-06	4.16E-05	6.91	4.91	-4.0
201812_s_at	hypothetical protein LOC201725 | translocase of outer mitochondrial membrane 7 homolog (yeast)	LOC201725 | TOMM7	201725 | 54543 4	.49E-06	5.50E-05	11.90	9.90	-4.0
228613_at	RAB11 family interacting protein 3 (class II)	RAB11FIP3	9727	1.62E-06	4.16E-05	7.59	9.58	4.0
228490_at	abhydrolase domain containing 2	ABHD2	11057	1.62E-06	4.16E-05	6.11	8.10	4.0
201617_x_at cal	desmon 1	CALD1	800	1.62E-06	4.16E-05	9.89	7.90	-4.0
218178_s_at	chromatin modifying protein 1B CH	MP1B	57132	1.62E-06	4.16E-05	9.14	7.15	-4.0
208050_s_at	caspase 2, apoptosis-related cysteine peptidase (neural precursor cell expressed, developmentally down-regulated 2)	CASP2	835	1.62E-06	4.16E-05	7.41	5.42	-4.0
1557938_s_at	polymerase I and transcript release factor	PTRF	284119	1.62E-06	4.16E-05	9.33	7.34	-4.0
214720_x_at s	eptin 10	10-Sep	151011	1.62E-06	4.16E-05	8.03	6.04	-4.0
206695_x_at	zinc finger protein 43	ZNF43	7594 1	.62E-06	4.16E-05	7.51	5.52	-4.0
217764_s_at	RAB31, member RAS oncogene family	RAB31	11031	2.71E-06	4.39E-05	9.25	11.24	4.0
200041_s_at	HLA-B associated transcript 1	BAT1 7	919	1.62E-06	4.16E-05	10.45	8.46	-4.0
205732_s_at	nuclear receptor coactivator 2 N	COA2	10499	1.93E-06	4.16E-05	4.87	2.88	-4.0

II.4 Clinical Database Analysis

Early diagnosis is imperative for the survival of victims of breast cancer. It is apparent that physicians need a method to predict probability of breast cancer to aid in early detection and prevention of the disease. Thus far, clinical data has not been used to predict the likelihood of breast cancer. The objective of this research task was to determine the relationship of controllable lifestyle factors such as alcohol consumption, exercise frequency, and medication to a woman's probability of developing breast cancer. If a methodology could be adopted that predicts the probability of a woman acquiring breast cancer, earlier measures could be taken for prevention and treatment of the disease.

This task utilized data collected via questionnaires on approximately 2,400 patients by WRAMC. The questionnaires consisted of approximately 245 questions, some with multiple parts, resulting in approximately 400 data items regarding medical history, genetics, and personal habits, such as smoking, drinking, and exercise. An extensive effort was required to condition the database in order to initiate statistical analysis, see Appendix 5.2. The research team originally received the clinical database questionnaire answers and corresponding Pathology Reports in a text file that had to be converted into an Excel file for analysis. Pathology data was then matched up with the respective

patients through their CBCP numbers, and then renumbered for analytical consistency purposes. Many patients had multiple pathology reports. The most recent report within 90 days of the questionnaire date was mapped to the patient's questionnaire data. Diagnoses from the pathology reports were divided into the categories of cancer, cancer related, and non cancer. For statistical significance, diagnoses with twenty patients or less were discarded.

Multiple choice questions were renumbered and reworded to form yes/no questions so that a binary code could be developed for every patient. Depending on whether a patient answered yes or no to a particular question, a binary number of 1 or 0 was inserted. The patient's binary code (ex.10010011101) yielded coherent information for every patient. Further details of the data conditioning approach are provided in Appendix 5.2.

Statistical analysis was performed on the conditioned data to correlate certain parameters within the database to the probability of a patient developing breast cancer. For each questionnaire data item, the percentage of patients with cancer, the percentage without cancer, and the percentage with a cancer-related condition were calculated. The differences between these respective sets were used to identify questions that could be sensitive to the probability of cancer. Parameters showing a high degree of correlation were noted and mapped back to relevant questions concerning lifestyle factors that could later help predict the probability of breast cancer. A more detailed description of the data conditioning and data analysis are reported in [2]. The conclusion of the analysis was that key lifestyle effects such as diet were not included in the questionnaire; therefore, the data could not be used to correlate all likely lifestyle causes of disease. Recommendations on changes in the questionnaire will be presented for review for inclusion in the revised questionnaire. The SMDC recommends a follow-on effort to correlate lifestyle parameters (i.e., does shaving under the arms relate to other lifestyle parameters) and use SMDC algorithms to predict likelihood of the patient getting breast cancer. Several of these parameters and other parameters which were expected to be important are discussed below.

Age

Increasing age is the most important risk factor for breast cancer in women. Breast cancer risk may be higher or lower depending on a woman's personal risk factors and experiences. Currently, a woman living in the US has a 12.3% (1 in 8) lifetime risk of developing breast cancer. In the 1970s, the lifetime risk of being diagnosed with breast cancer was 9.1% (1 in 11). This increase is due to longer life expectancy, as well as increases in breast cancer incidence due in part to long-term hormone replacement therapy and the rising prevalence of obesity. The WRAMC database results were: age > 60 = a strong trend toward cancer, age 51-60 = a neutral trend, age 41-50 = a trend toward non cancer, and age 31-40 = a strong trend toward non cancer.

Family History

The WRAMC database did not yield trends relating to family history. However, the American Cancer Society has shown that women with a family history of breast cancer, especially in a first-degree relative (mother, sister, or daughter), have an increased risk of

developing breast cancer. The risk is higher if more than one first-degree relative has developed breast cancer. The risk increases the younger the relative was at the time of diagnosis. It is estimated that 5% to 10% of breast cancer cases result from inherited genetic mutations or alterations in the breast cancer susceptibility genes. These genes have been identified as BRCA1 and BRCA2. These mutations occur in far less than 1% of the general population. From population-based studies, women with BRCA1 mutations are estimated to have a 65% risk for developing breast cancer by age 70. The corresponding risk for BRCA2 mutations is 45%. Scientists believe that most of the occurrence of breast cancer in families results from the interaction between lifestyle factors and low risk variations in genetic susceptibility that may be shared by women within a family. Given the apparent disagreement between the WRAMC data and the American Cancer Society, the data is neutral. This result can be due to the interpretation of the questions asked or the data conditioning needed to process the results.

Hormonal Factors

Reproductive hormones are thought to influence breast cancer risk through effects on cell proliferation and DNA damage, as well as promotion of cancer growth. Early first menstrual period (<12 years of age), older age at menopause (>55 years of age), older age at first full-term pregnancy (>30 years of age), and fewer number of pregnancies may increase a woman's risk of breast cancer by affecting the levels of reproductive hormones produced by her body. Breastfeeding has consistently been shown to decrease a woman's risk of breast cancer with greater benefit associated with longer duration. Recent use of some oral contraceptives may slightly increase the risk of breast cancer; however, women who have stopped using oral contraceptives for 10 years or more have the same risk as women who have never used the pill. The database shows that use of combination hormone replacement therapy (HRT), which combines estrogen and progestin, resulted in an increase breast cancer risk, with higher risk associated with longer use.

Obesity

Using the WRAMC database the following "obesity" trends were noted. BMI<20 (under weight) = neutral, BMI 20-30 (normal) = slight cancer group trend, BMI 30-40 (obese) = slight trend toward cancer group, BMI > 40 (morbidly obese) = neutral. Another recent study found that women who gained 55 pounds or more after age 18 had almost 1.5 times the risk of breast cancer compared with those who maintained their weight. A gain of 22 pounds or more after menopause was associated with an increased risk of 18%, whereas losing at least 22 pounds after menopause and maintaining the weight loss reduced risk. In postmenopausal women, circulating estrogen is primarily produced in fat tissue. Thus, having more fat tissue increases estrogen levels and the likelihood of developing breast cancer. A measure of obesity is indicated by calculating the body mass index (BMI) The BMI (in US units) = weight (lbs) x 703 / height2 (in2).

Physical Activity

Growing evidence supports physical activity as having a small effect on breast cancer susceptibility. Although most studies find reduced risk in women who exercise vigorously for 45 to 60 minutes on 5 or more days per week. One study suggests that regular physical activity, regardless of intensity, may reduce the risk of breast cancer in

postmenopausal women. The WRAMC data indicated that women who exercised three times a week for twenty minutes or less had marginal improvement in risk of breast cancer. And those women who exercise more than thirty minutes three times a week were less susceptible to breast cancer, see Figure 3.

		Non-Cancer Group		Cancer Group	Cancer Related Group
1 time per week	20-30 mins	21%		15.8%	18.9%
	30 mins or more	35.3%	11.6%	23.7%	23.8%
1-3 times per week	20-30 mins	10.6%		10.8%	8.4%
	30 mins or more	61.5%	17.5%	44%	45.5%

Figure 3: Exercise Results

Alcohol

The WRAMC data was obtained for women in age brackets and varying consumption. In the 66 to 75 age bracket, for women who never drank the susceptibility to breast cancer was slightly higher. The same trend existed for women in that age bracket who drank frequently. This result is probably masked by age susceptibility. For pre-menopausal women in the 36 to 45 bracket, those who never drank had a slight susceptibility while those who drank frequently had a slight trend toward non-cancer. Overall this data was non-conclusive. Other studies indicate that alcohol consumption is consistently associated with increased breast cancer risk. Analysis of more than 40 studies suggests that the equivalent of 2 drinks a day (or 24g of alcohol) may increase breast cancer risk by 21%. This increased risk is dose-dependent, and exists regardless of the type of alcoholic beverage consumed. A recent review concluded that the most likely mechanism by which alcohol increases risk of breast cancer is by increasing estrogen levels. Thus, reducing alcohol intake may be a useful strategy for reducing breast cancer risk among regular consumers of alcohol.

Tobacco

All of the WRAMC data associated with smoking was indeterminate. There was very little difference in cancer and non cancer patients over years of smoking, number of packs smoked, and smoke free. Some other studies have found no link between active cigarette smoking and breast cancer. Though both active smoking and secondhand smoke have been suggested to increase the risk of breast cancer in a number of studies that

restrict the comparison group to women who report no exposure to secondhand smoke, this issue remains controversial.

Hormone Replacement Therapy

The WRAMC data indicated a strong susceptibility to breast cancer from hormone replacement therapy in general. Estrogen replacement alone had a neutral effect. However, Estrogen and Progesterone in combination produced a strong susceptibility. Use of combined hormone replacement therapy, estrogen and progestin therapy, increases the risk of breast cancer, as well as the likelihood that cancer will be found at a more advanced stage. Hormone replacement therapy may increase the risk of late-stage diagnoses by increasing breast tissue density, thereby reducing the effectiveness of mammograms.

Chemoprevention

The WRAMC data indicates that women who take tamoxifen as prescribed have a significantly less likelihood of getting breast cancer. Several other clinical studies have shown that, in women known to be at increased risk for breast cancer, the drugs tamoxifen and raloxifene may reduce this risk. Tamoxifen is currently used for the treatment of both early and advanced breast cancer in pre- and post-menopausal women. It is also approved for the prevention of breast cancer in women at high risk of developing the disease. It has been further approved for the reduction of cancer in the opposite breast. After an average of 7 years of follow up, breast cancer risk decreased by 42% in the group that received tamoxifen. These long-term, follow up results indicate that the reduction in risk persists after completion of the 5-year treatment schedule. A study comparing the effectiveness of the two drugs, called the Study of Tamoxifen and Raloxifene (STAR) trial, found that raloxifene reduced the risk of invasive breast cancer to the same degree as tamoxifen.

Other Data

Many of the above results corresponded to the American Cancer Society research information. However, some results are either contradictory or seem to defy common sense, such as:

1) Women who shave under their arms have a significantly reduced susceptibility to breast cancer.
2) Women who wear an under wire bra have a significantly reduced susceptibility to breast cancer.
3) Women who use deodorant have a significantly reduced susceptibility to breast cancer.

Clinical Data Conclusions

Multi-variant effects are not analyzable using data that is meant to be viewed by a human. Examples of these multi-variants are exercise and breast density, density from cancer and density from age, and body mass index and exercise.

If a questionnaire is to be interpreted visually and subjectively by qualified personnel who are familiar with the data and no statistical analysis is needed, then the questionnaire can be similar to the one used at WRAMC, Clinical Breast Care Project. However, if analysis of the results and possible interrelationships between the factors is a desire, then the questionnaire must be carefully crafted to yield data that is analyzable.

Consideration must be given to significant sample size, education, age, question relevance, and conversion of all data to a format this is recognizable by a computer and the appropriate software. Even though the number of patients in the database was quite large, most were referred due to suspicion of cancer. This may skew the results when compared to a normal population.

II.5 Mammogram Image Processing

The occurrence of breast cancer in the female population in the United States has been on an ever increasing trend as seen from American Cancer Society reports. Early detection and treatment are the keys to survival. The detection of cancerous tumors in the presence of dense tissue and other artifacts has many similarities to the detection of hostile targets in clutter and false targets using radars, infrared sensors, and sonar for the Department of Defense. In these military systems, detection rates of better than 95% are required. The ACRIN Digital Mammographic Imaging Screening Trial (DMIST) comparing breast cancer detection using digital mammograms to film mammograms found that only 70% of cancer was detected using either film or digital mammogram data in the normal population. Digital mammograms were better at detecting cancer in dense breast: 70% were detected with digital mammograms and 55% with film mammograms. The comparatively poor detection rates of mammograms suggest a need to provide the radiologist with better tools to detect cancer. The SMDC proposes the use of algorithms and techniques used in missile defense to improve the probability of detection of cancer. Also, SMDC proposes to use breast density change algorithms to queue the radiologist to areas in the breast with high probability of cancer. The SMDC obtained an ACRIN digital mammogram data set for this study consisting of 11,528 mammogram images from 2,467 patients. 1,503 of the images were from the 305 patients who tested positive for breast cancer.

SMDC, Decibel Research, and the University of Alabama Huntsville (UAH) have developed image processing algorithms to detect and discriminate targets in clutter. These algorithms were developed and tested using University of South Florida (USF) digitized film mammogram database, Walter Reed Army Medical Center (WRAMC) digitized film mammograms, and ACRIN digital mammograms. Most recent efforts have been concentrated on ACRIN digital mammogram image data set consisting of 11,528 mammogram images from 2,467 patients. 1,503 of the images were from the 305 patients who tested positive for breast cancer.

Conditioning of the ACRIN database for processing and analysis was a major effort. It consists of digital mammograms from four vendors. The dynamic range of the intensity

data varied from vendor to vendor; therefore the data has to be processed independent of the contrast stretching used. Also, many of the images were inverted and had to be converted to the negative format. There were artifacts in the images that had to be removed. Conditioning of the data consisted of the following tasks: 1) Isolate the breast from other artifacts in the image and define a bounding box for the breast; 2) Indentify and correct inverted images; 3) Choose a test database consisting of cancer negative and cancer positive patients; 4) Divide the breasts into low density, medium density and high density categories; and 5) Identify the location of cancer in positive patients.

The SMDC approach for processing mammogram imagery to improve the probability of early detection and discrimination of breast cancer is shown Figure 1. The approach is to use changes in breast tissue density across the breast, and differences in density between the left and right breast to warn of changes in the breast and to queue areas that are different. Once these areas are identified the ADA detection algorithm is executed to identify anomalies in the area of interest (or the radiologist may investigate the area of interest). Next the anomaly discrimination algorithms are executed to discriminate the anomalies as cancerous or non-cancerous. If there is a high probability that the anomaly is cancerous the ADA algorithm can be executed in a higher resolution mode to better define the margins.

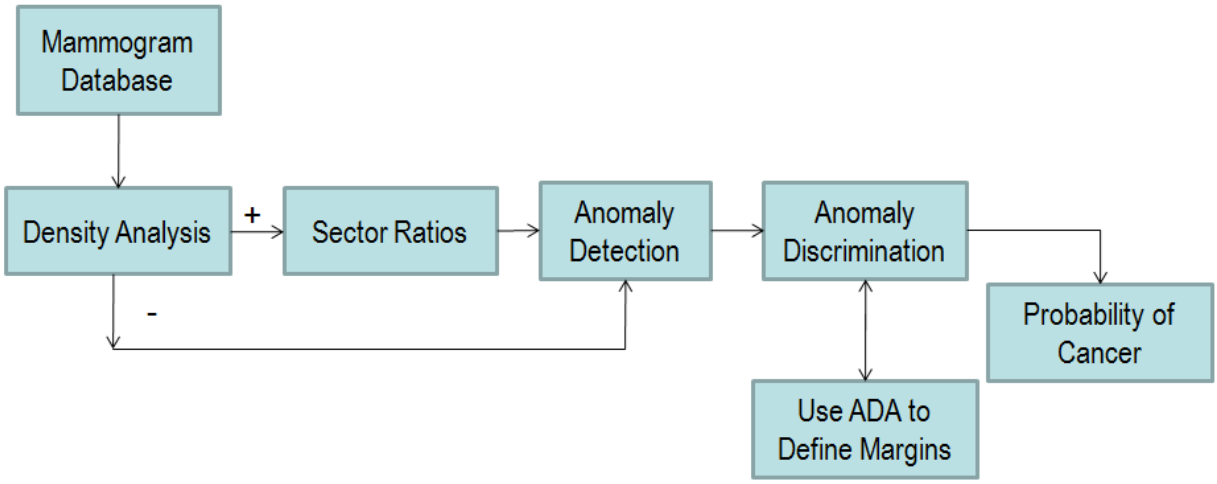


Figure 1. Mammogram Image Processing Approach

Breast Density Analysis

SMDC is investigating early detection/prediction of breast cancer based on changes in breast density. Calculation of the absolute change in breast density on digitized film data year to year cannot be performed because the film exposures are different year to year. An example of WRAMC provided digitized film data on a patient where the mammograms were collected over multiple years is shown in Figure 5. In this case we show digitized mammograms for four years. The film exposure and orientation of the breast is different year to year. Due to lack of calibration the change in density across the breast year to year was the only reasonable way to perform the analysis. Since absolute

breast density cannot be measured accurately using film mammograms, SMDC developed algorithms to calculate changes in breast density across the breast and to compare density in the left breast to density in the right breast.

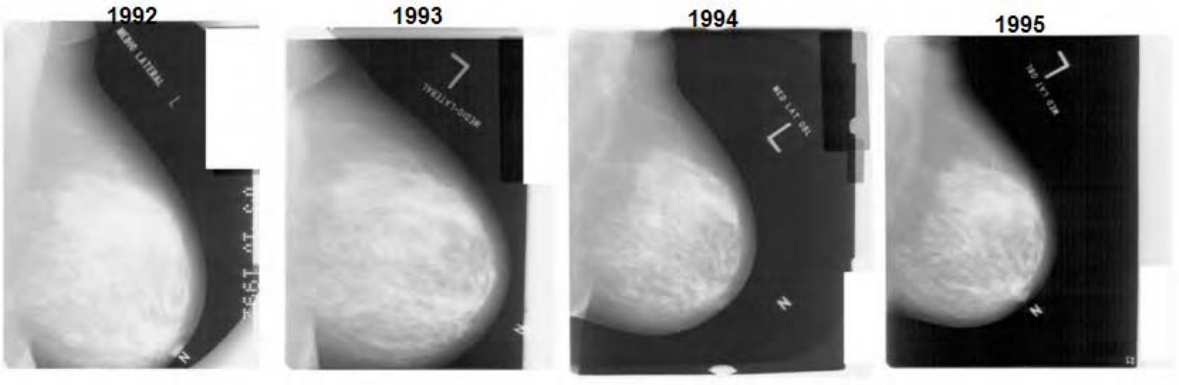


Figure 5 Breast Density Change Detection Year-to-Year Image Registration Issues

The change in density across the breast was calculated by partitioning the image into boxes, see Figure 6.

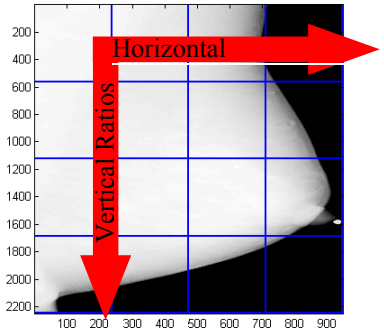


Figure 6. Breast Image Partitioning

Ratios of the average intensity of the boxes were calculated by calculating the mean intensity of each box and then calculating the ratios of the boxes horizontally and vertically, see Figure 7.

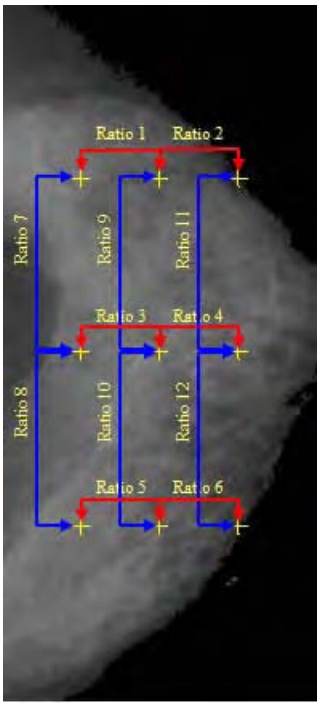


Figure 7. Partitioned Breast Ratios

In this case the breast image was partitioned into 9 boxes resulting in 6 horizontal ratios and 6 vertical ratios. These ratios were calculated on mammograms of a cancer patient that we had images seven years from 1992 to 2001 when the cancer was detected. The plot of the ratios versus ratio numbers for the different years is shown in Figure 8.

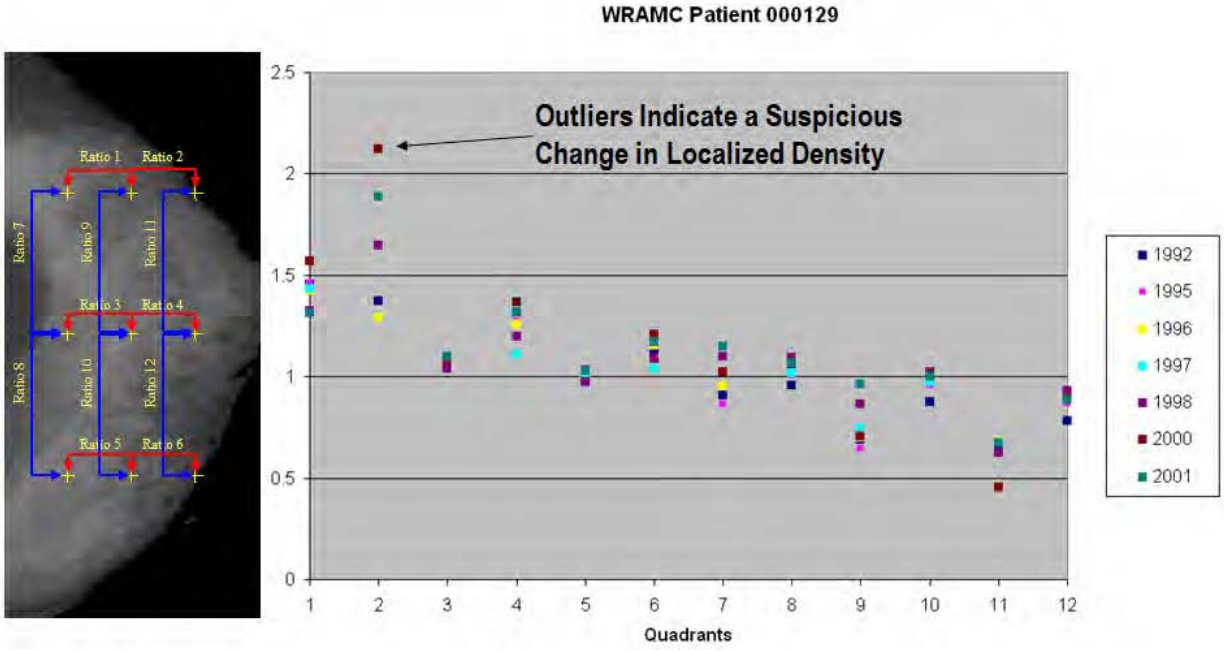


Figure 8. Ratios of Intensities across the Breast

In this case there is a distinct change in the area of ratio 2 starting in 1998. This result was then further checked by calculating the mean intensities in the cancer positive and cancer negative breasts and taking the difference. The results were the same. The differences in the mean intensities started to increase dramatically in 1998 and continued to increase until 2001 when the cancer was detected. These results suggest that the onset of cancer or a pre-cancerous condition could have been detected in 1998 three years before the cancer was detected. Also by using the ratios the location of the change was determined. These results led to development of more sophisticated breast density change detection algorithms. These algorithms were executed on the ACRIN digital mammogram database.

Digital mammograms in the ACRIN data base were well calibrated; therefore changes in the density can be accurately calculated. This led to an in depth study to determine if slight changes in density due to a cancerous anomaly could possibly be detected by comparing the density of the left breast to the density of the right breast. Several techniques were implemented to identify and classify this density change. We used the intensity of the images to directly relate to the density of the breast.

The SMDC's initial study of the ACRIN data compared the average intensity in the left breast to the average intensity in the right breast. The boundaries found in the data conditioning process allowed only the intensity in the breast to be considered. This enhanced the accuracy of our study. To further enhance this process, we partitioned the breast image into blocks and analyzed the average intensities in each individual block as described above. A study was conducted to identify the optimum block size. It was discovered a 4x4 grid provided the best results, see Figure 9.

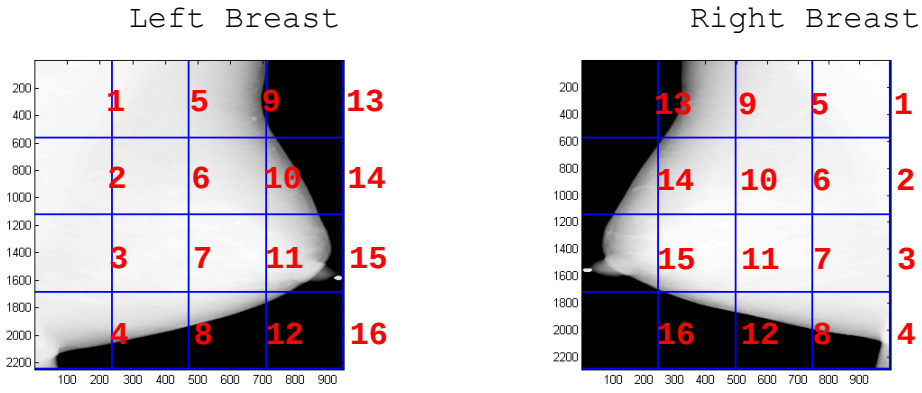


Figure 9. Partitioning the Breast Image into Blocks

Various statistical algorithms were applied to these blocked mammograms. One of the methods used to detect the region of interest was comparing the change in intensity

across the breast. Ratios of the blocks in the horizontal and vertical direction were calculated.

These ratios were then compared in the left and right breasts. An algorithm used the ratios in the left and right breasts to detect the blocks with the greatest changes from left to right and vice versa. This algorithm categorized the blocks of interest by labeling them in descending order based on greatest change. The results of using this approach are shown in Figure 10. In this case the maximum ratios are in the breast with cancer. The blue region on the breast is where the cancer is located.

Patient 1084
MLO View
Ductal Cancer in Left Breast @ 11-12 Anterior

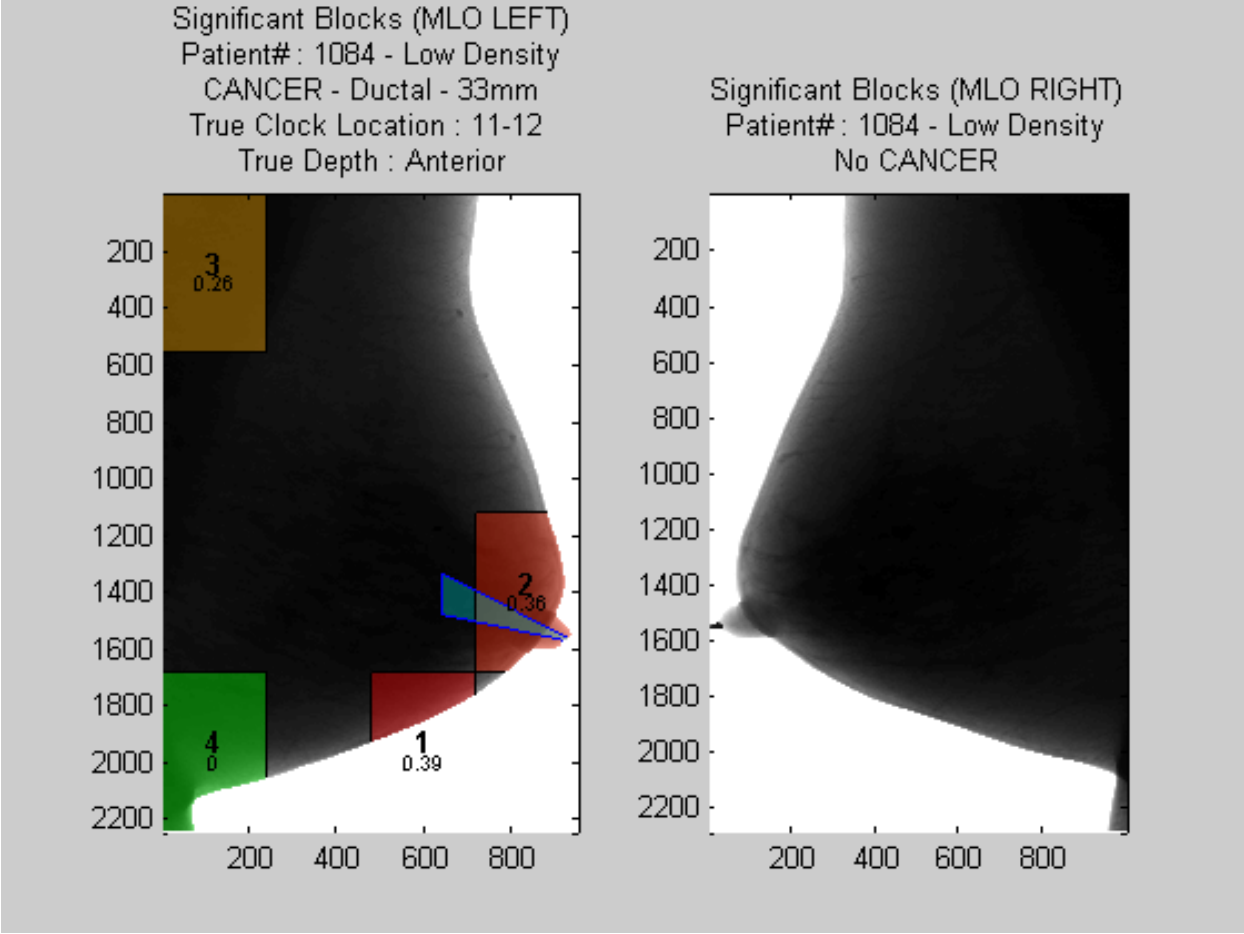


Figure 10. Block-to-Block Intensity Ratio Comparison for Left and Right Breast

The SMDC has developed algorithms to compare difference in intensities from block to block which is also providing promising results. Since the digital mammogram data is so well calibrated year to year comparisons of density should be accurate. This will allow for other image processing techniques to be applied such as year to year mammogram image subtraction.

Breast Density Analysis Conclusions

The results from analysis SMDC performed on breast density change across the breast and differences in density from left to right breast indicate a probability that these methods can be used for detection of pre-cancerous condition and early detection of cancer. They can also be used to cue radiologists to areas of interest.

Digital mammograms well calibrated and are easily stored and processed supporting more sophisticated year to year analysis. Recommend further development of image processing algorithms to perform year to year processing.

Anomaly Detection

Anomaly detection is the process of isolating discrete objects from the cluttered background of the breast tissue. The overall image processing workflow consists of (1) segmentation of mammogram image to crop the image down to just the region of interest (ROI) occupied by the breast tissue (see Figure XX), (2) detection of masses in the breast tissue, (3) classification of detected masses as cancerous or non-cancerous, and (4) the definition of margins of cancerous masses to support surgical procedures.

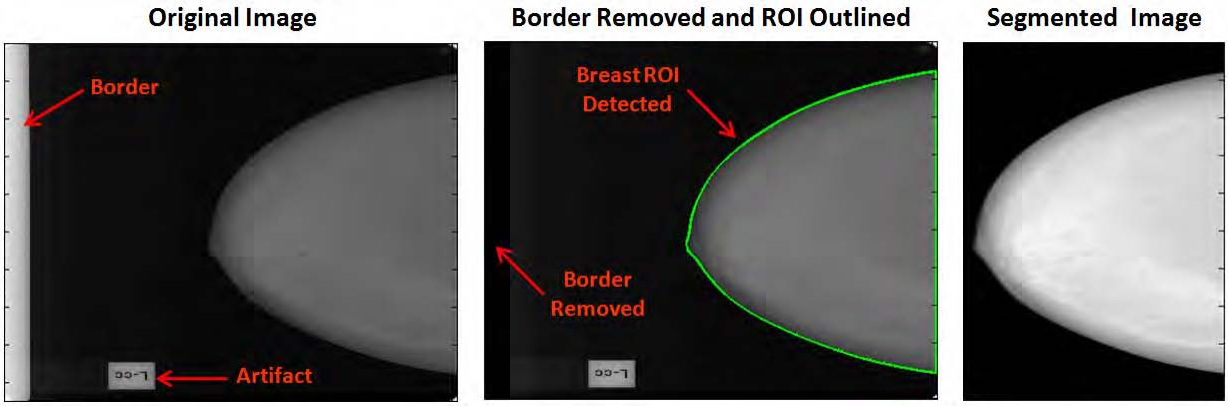


Figure XX. Image Segmentation Example

This project examined two techniques for anomaly detection: the Anomaly Detection Algorithm (ADA) and Pulse-Couple Neural Networks (PCNN). Both of these algorithms operate in similar ways to perform object detection.

The Anomaly Detection Algorithm (ADA) identifies objects in an image, even when those are objects are in a cluttered and/or noisy environment. The ADA works on the principle of threshold detection, enhanced by the sharing of intensity information among neighboring pixels. Throughout the execution of the algorithm, a global threshold is slowly relaxed, and newly detected pixels adjacent to an anomaly are associated with that anomaly. Detection of a single, isolated pixel, can lead to a cascade of detection by the pixel's neighbors, based on the original pixel linking its intensity energy to its neighbors. An example image showing ADA's ability to identify objects is shown below in Figure XX. Note how the muscles are clearly identified as objects.

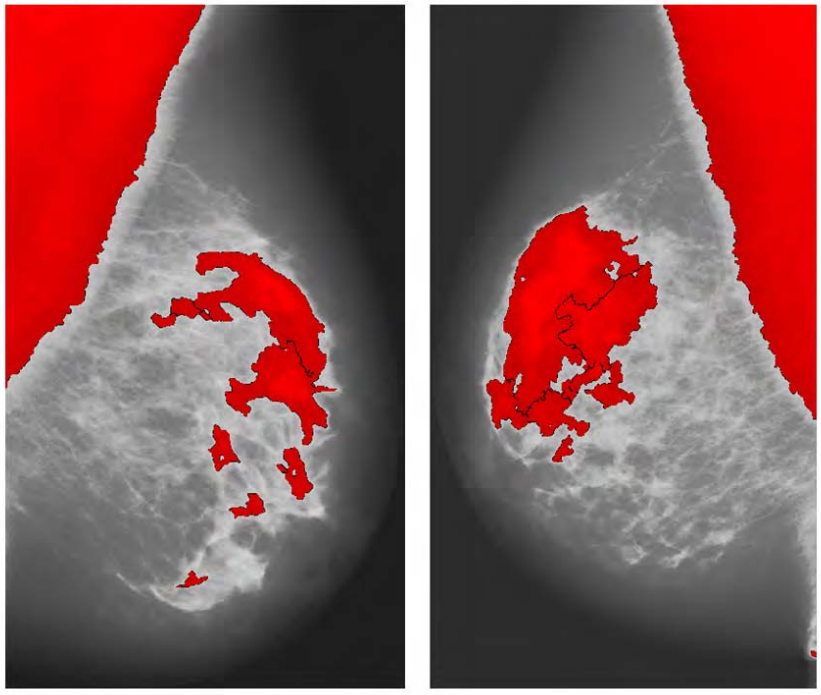


Figure XX. ADA Detections in ACRIN Digital Mammograms

PCNN is a biologically inspired algorithm based on Eckhorn's model of the cat visual cortex. It has been shown that PCNN are capable of image segmentation, smoothing, feature extraction, noise reduction, etc [1,2]. Figure XX below shows an example of PCNN detections. PCNN works well at identifying regions of similar pixels, enabling it to detect individual objects such as cancerous lesions.

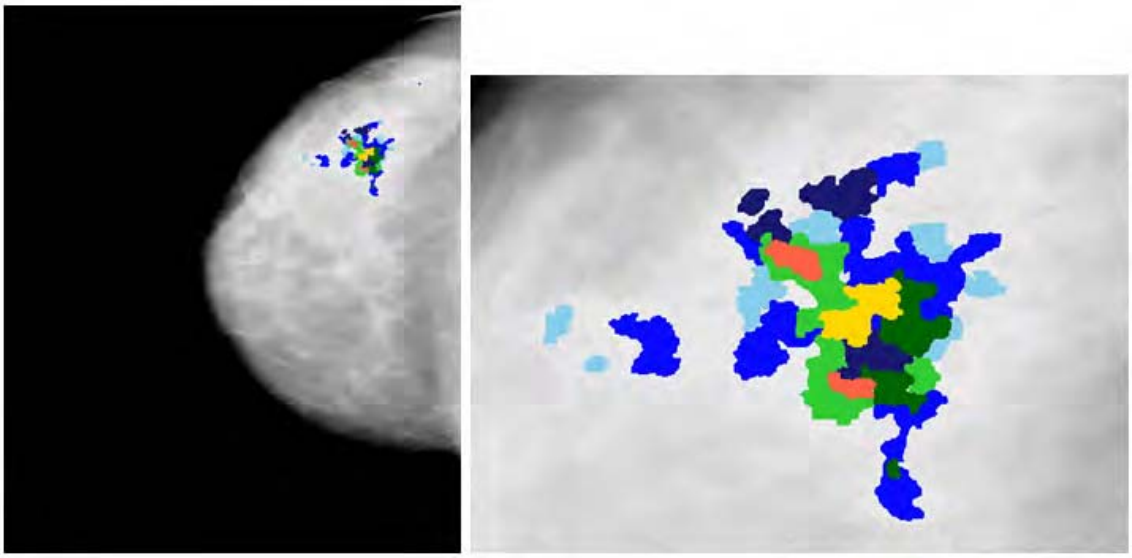


Figure XX. PCNN Processed Image

Anomaly Classification

The SMDC developed algorithms to extract geometric classification features of anomalies such as distribution of radii from the center of energy and changes in the radii to describe the variation in the margins of the anomalies, see Figure 11.

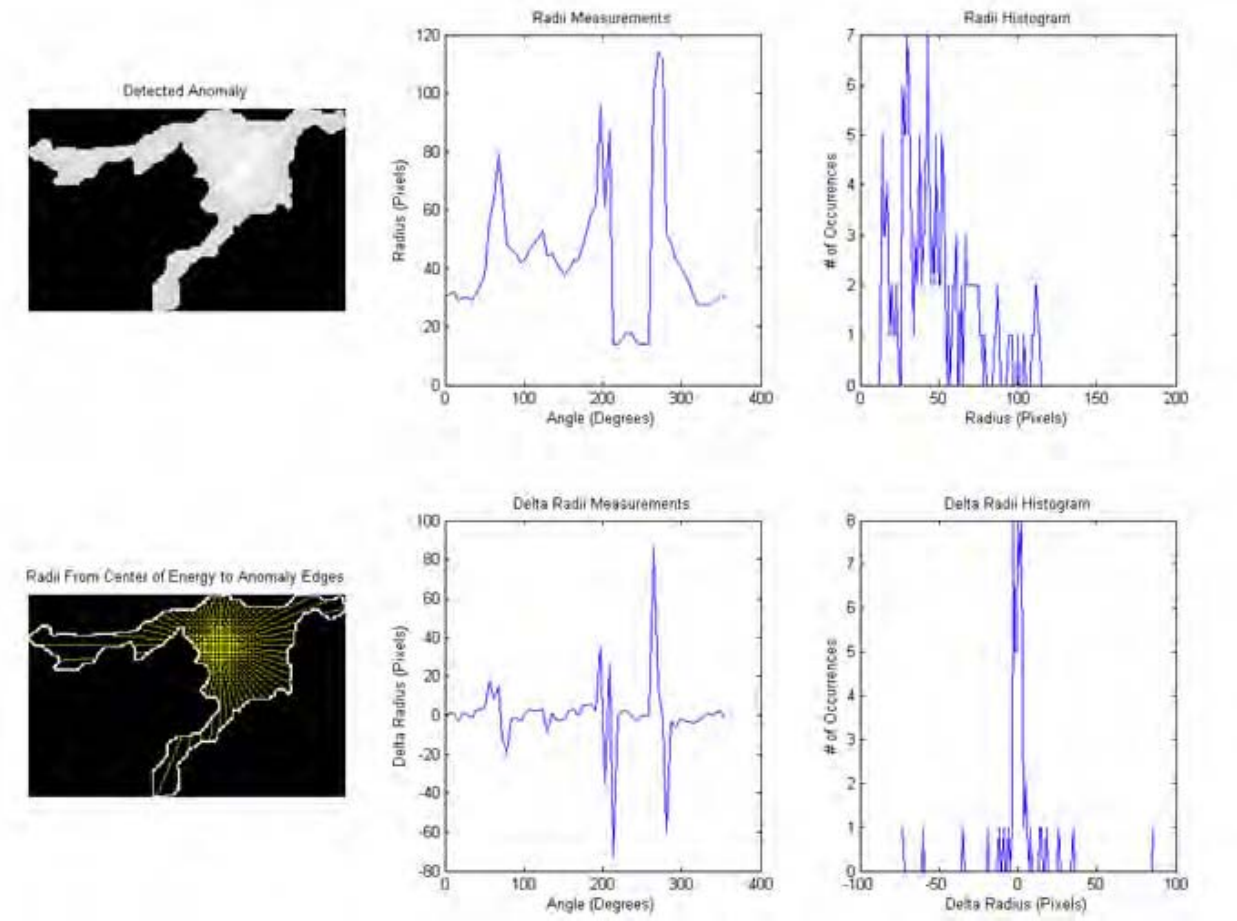


Figure 11. Anomaly Classification Features

These features and features utilizing spectral analysis are used to discriminate intensity-only images anomalies (such as digitized film mammograms, digital mammograms, and ultrasound images) . Spectral analysis is performed by developing discrete sequences that can be analyzed by well-established techniques as well as emerging digital signal processing (DSP). One implemented method of sequence generation is to index intensities by the radial distance of pixels from a center of energy, see Figure 12. This profile is for a cancerous anomaly and distinct from other non-cancerous anomalies.

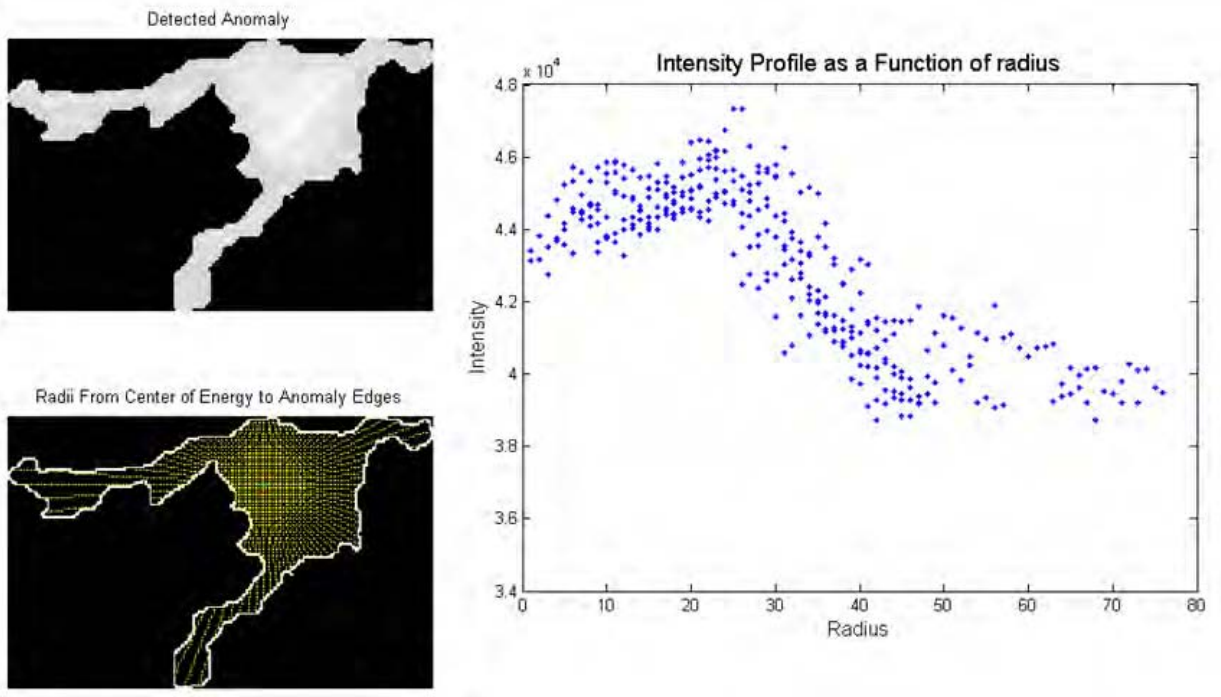


Figure 12. Intensity Profile of Radii Discriminant

The basic DSP technique is the decomposition of a sequence into orthogonal harmonic components using the Fourier transform. The Fourier transform preserves the inherent symmetry of the magnitude data as Hermitian symmetry. SMDC examining the asymmetric complex Fourier coefficients for robust exploitable features associated with true positives. This is the simplest of many potentially applicable DSP techniques. The results from performing the spectral analysis These results show that each anomaly has a discrete signature. Cancerous tissue should have discrete signatures and there is a possibility that different types of cancers will have different enough signatures to be discriminated from other non-cancerous anomalies.

on three anomalies are shown in Figure 13.

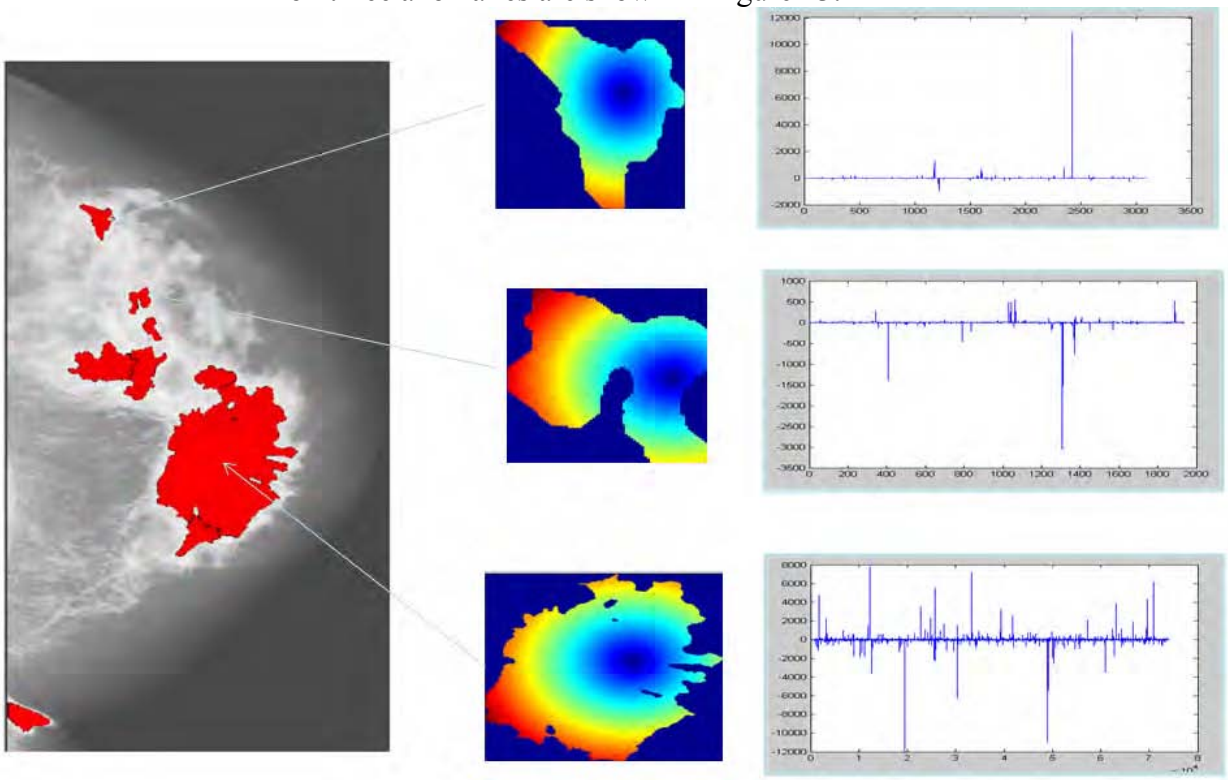


Figure 13 Anomaly Spectral Analysis Example

The features of the anomaly derived using the techniques described above can be used in n algorithm such as SMDC's Genetic Response Surface Model to predict the probability that anomalies are cancerous.

Anomaly Classification Conclusions

The algorithms developed to extract features of anomalies provide distinct signatures that can be used to discriminate cancer from other anomalies.

Recommend that these algorithms be executed on both cancerous and non-cancerous anomalies for validation.

Recommend that the algorithms be integrated with the GRSM to provide anomaly discrimination architecture.

Ultrasound Image Processing:
Theoretical and experimental results based on the laws of propagation of ultrasound waves in biological tissues are available. Acoustic speed, rate of wave attenuation, and acoustic modulus of elasticity in ultrasound waveforms have been used for determining masses of biological tissues. Changes in echo waveforms caused by microscopic variations in the mechanical properties of tissue can reveal complementary information to mammogram images. Ultrasound waveforms provides additional component to feature-

based discrimination as time-frequency representations and time-series analysis methods readily apply to the analysis of the waveforms.

The SMDC did not have an ultrasound data base with truth data sufficient to develop discrimination algorithms. Ther is a problem with registering ultrasound images with other images such as mammograms. SMDC believes that the method for collecting ultrasound data will have to be changed to image a larger portion of the breast before an automated process for using the data can be developed.

III. Key research accomplishments

- This project started with an excellent vision, but in execution we found a main burden in data collection. There are a total of only 45 subjects meeting the original subject selection criteria, which truly reflects the difficulty of multidisciplinary study especially when multiple clinical imaging platforms are needed. We found alternative approaches to maximize the value of the outcome of this project.
- SMDC processed data collected via questionnaires on approximately 2,400 patients by WRAMC. The questionnaires consisted of approximately 245 questions, some with multiple parts, resulting in approximately 400 data items regarding medical history, genetics, and personal habits, such as smoking, drinking, and exercise. An extensive effort was required to condition the database in order to allow for computer-aided processing. SMDC performed an analysis correlating lifestyle and family history to the likelihood of cancer, and compared these results to other clinical studies.
- SMDC developed and tested image processing techniques for early detection of pre-cancerous and cancerous lesions in the breast.
- SMDC developed a methodology to classify anomalies in the breast as cancerous or benign.
- Gene expression microarray experiments are performed on blood samples from 92 subjects and breast tissues from 37 subjects.
- In both the experiments using blood samples and tissue samples, we found differentially expressed gene patterns. Using Wilcox test and FDR control, we found about 400 gene differentially expressed in blood samples between normal and cancer groups.
- In microarray data using tissue samples, we found about 6000 genes changed between normal and cancer group and 1400 of them with FC greater than 3.
- From the clustering analysis, correlation analysis, and PCA results, we found the difference and patterns between the gene expression profiles between normal and cancer subject.
- Comparing the gene expression profiles between blood samples and tissue samples, we found that the expression profiles between the tissues are more distinct than those of blood samples between normal and cancer group.

IV. Reportable Outcomes

None.

V. Conclusions

The clinical questionnaire is not designed for easy use in computer-aided analysis. The database needs to include information on diet and other lifestyle factors if it is to be used to predict the likelihood of disease.

The SMDC image processing algorithms show promise at detecting and classifying masses in the breast. These algorithms require additional testing on the entire ACRIN digital mammogram data set. The algorithms need to be integrated into an architecture to be used to predict pre-cancerous conditions and to detect / discriminate cancer in the breast.

Gene expression data from breast tissues can accurately separate invasive cancer subjects from benign disease subjects. But data from blood samples cannot readily do it. Note that regarding the ways specimens are obtained, breast tissue samples requires surgery but drawing blood samples is minimally invasive. We see potential scientific value of the gene expression studies executed here, and we will further work on the analysis of the data to reach a solid conclusion and possibly a publication on this topic.

When more testing has been completed by SMDC in applying the missile detection technology to breast cancer detection, merging of the molecular findings with image analysis technologies can potentially improve the detection of breast cancer. This could result in an integrated CBCP systems architecture for pre-cancerous and cancerous diagnoses. Tina please modify this based on the SMDC portion.

VI. References

See report above

VII. Appendices

See report above